NSPW Responses to Agency Comments
Baseline Ecological Risk Assessment
Draft Remedial Investigation Report
Ashland/NSP Lakefront Superfund Site

EPA Region 5 Records Ctr.
313799

COMMENTS ON BERA

General Comments

1. It is our understanding that the BERA would incorporate data collected for the 1998 and 2001 ERAs performed by SEH. While SEH data was used to select COPCs and to propose sediment cleanup levels, it appears that this data was not incorporated into the food chain modeling. Fish tissue from SEH study should also be included.

Response

We also used SEH data from the 2001 bioassays. The SEH fish data was not used because the data consisted of 19 fish of 10 species caught at two times of the year, 5/98 and 10/98. Of these fish only nine had detectable levels of PAHs. The PAHs detected in these nine fish were all low molecular weight PAHs. There were no data from fish caught at reference locations. In addition, PAH levels in these fish were, with only a couple of exceptions, substantially less than they were in the fish caught during the RI sampling.

In NSPW's opinion this results in a poor database to use as the basis for any food chain modeling. We request the reviewer's explain the rationale for using these data.

2. We agree that site sediments should be addressed in the FS. However, the impacts from soil and sediments to higher level organisms has not been adequately characterized in the BERA, so conclusions on excluding these pathways is premature at this time.

Response

We do not agree. Impacts to higher level organisms have been addressed and the conclusions of this evaluation are consistent with results of evaluations of the risk to higher level organisms at other MGP sediment sites: wildlife are not at risk.

The results of the BERA indicate that risks to higher level organisms exposed to PAHs in soil and sediment are negligible. Food web models presented in the BERA indicated no HQs greater than 1.0 for wildlife exposed to PAHs in soil or sediment. These models were based on conservative estimates of PAH accumulation in prey items: USEPA Eco-SSL bioaccumulation models for terrestrial receptors, site-specific measurements of PAH concentrations in fish, and estimated concentrations of PAHs in benthic invertebrate based on the results of the site-specific Lumbriculus variegatus bioaccumulation study. The absence of adverse effects in higher level organisms is corroborated by Eco-SSL guidance, which indicates that the bioaccumulation of PAHs is expected to be minimal due to the rapid metabolism of these compounds after ingestion by birds and mammals (USEPA 2005). Therefore, prey items of higher level organisms are not expected to contain concentrations of PAHs that would cause adverse effects to upper trophic receptors. Based on these results, we believe that impacts to higher level organisms have been adequately characterized in the BERA as negligible.

3. The BERA does not appear to address the free product in the bay area. Therefore, it is assumed that removal of the free product will be addressed in the RAOs and FS.

Response

The potential for adverse effects to ecological receptors from releases of contaminants from subsurface sediments (including contaminants in sheens) is discussed in Section 6.2.14 of the BERA. In addition, the effect of "free product" releases as a source to sediment is addressed indirectly since Exposure Point Concentrations (EPCs) for sediment were based upon measurement of VOCs and SVOCs in that sediment. Any release of "free product" that impacted sediments at the Site would have been accounted for in these measurements.

The removal of free product will be addressed in the FS.

4. The shallow groundwater discharge to the bay area does not appear to have been addressed.

Response

The effect of the shallow groundwater discharge to the bay, if any, is incorporated into the EPCs for surface water and sediment. Any contaminants transported to the bay area, would need to become associated with bay surface water or sediment to be of potential risk to ecological receptors there.

5. We do not concur at this time that the relative significance of the lines of evidence presented in Section 4.3.2 (numbers 1 through 3) are appropriate for this site. Further characterization of site risks and the uncertainty associated with each line of evidence needs to be performed before relative significance is assigned to a line of evidence.

Response

This comment should refer to Section 4.2.3.

NSPW believes that the relative significance of these lines of evidence as the basis for assessing risk has been well-established in guidance and the literature for ecological risk assessment, i.e. that it is ultimately the population and community level response that is important. In that the benthic community assessment had sufficient discriminatory power to assess at a minimum moderate differences between Site and reference station benthic communities, we believe this line of evidence should be accorded the greatest weight.

An implicit discussion of uncertainty was provided for the line of evidence that involved comparison of site-specific media concentrations and/or estimated ingested contaminant dose estimates (the latter for wildlife) to effects levels [toxicological benchmarks and TRVs] for the various receptors of concern (ROCs) when we discussed that the effects levels are generic and not site-specific. Uncertainties associated with the sediment toxicity tests and the benthic community assessment were discussed in the specific reports of these lines of evidence Attachments 2 and 3 to

Appendix B. In addition in many places in Section 6.3 the uncertainties associated with these latter two lines of evidence are discussed.

However, in the final BERA we will provide a more explicit discussion of the relative uncertainty including an analysis of the discriminatory power of the various analysis used in the benthic community assessment.

6. Wood chips were commonly used in the purification process at MGP sites. Therefore, without documentation and proof that the wood chips from the purification process were not disposed of in the ravine and in the lower bluff and harbor area, it cannot be concluded that the wood chips are present solely due to non-MGP processes (lumber operation).

Response

As described in the response to General Comment 1 to the draft RI report, the MGP was a water gas facility throughout its history. Water gas plants did not produce significant amounts of purifier residue, as confirmed by the absence of nitrogen bearing compounds in the environmental samples collected. Additionally, although some wood residue was encountered during investigations in the ravine fill this material is not extensive, nor did it show indications that it was related to the MGP. There is no evidence supporting the supposition that wood in the bay is a result of purifier box waste from the MGP.

As indicated in Section 2.2, the presence of lumbering operations and log rafting is well documented and by itself could explain the presence of the wood found in Site sediments.

"Beginning in the mid to late 1800's, this area was filled with a variety of materials including slab wood, concrete, demolition debris, municipal and industrial wastes and earthen fill that created the land now occupied by the park. The filled area was used for lumbering and sawmill activities which occurred during the deforestation of the northern portion of Wisconsin around the turn of the century. Timber was also cut in various places in the area, including the Apostle Islands and the Arrowhead region of Minnesota, and the logs rafted into the Ashland area where they were floated awaiting processing. The large amount of wood "mulch"[1] in aquatic portions of the Site provides testimony to the log rafting that occurred here. The extensive amount of bark mulch-sized wood particles and small wood chips found on and in the sediment today likely originated from the constant working of the logs in the log rafts as well as, perhaps, the disposal of wood debris from the saw mill operation."

7. When presenting total PAH concentrations and normalized organic carbon (NOC) PAH concentrations, the associated organic carbon content should also be presented.

[1] The term wood mulch probably best describes the conditions of most wood debris and wood waste found overlying the surface sediment. Most of it is ground up pieces or bark and twigs not unlike bark mulch. In addition to the wood mulch there is a variety of other wood waste including logs, shingles and other manufacturing wood waste, branches, and twigs.

Response

Although we are unaware that this is a convention, and indeed, very few of the scientific journals focusing on environmental chemistry and risk follow this convention, we will include this information in the revised BERA.

8. The shallow soil exposure point concentrations (EPCs) used in the human health risk assessment (HHRA) for recreational exposure in Kreher Park differ from the EPCs used to evaluate soil exposure by the mouse and blackbird. Why do the datasets differ?

Response

The BERA excluded two soil samples which were used in the HHRA. In one instance it was because the soil interval was uncertain, in the other because it indicated it was a "seep boring". We will reconcile soil sample selection criteria with the HHRA so they are consistent.

The BERA also defaulted to 95%UCL based upon a bootstrap or a jackknife if the EPC data were not normally distributed, while the HHRA used the Chebyshev estimation of the 95% UCL that ProUCL recommended. This will be re-evaluated in the revised BERA.

9. Ecological RAOs presented in the RI/FS (Appendix A) will need to be adjusted after the BERA is corrected.

Response

We will re-evaluate the proposed RAOs once the issues raised in these comments are resolved.

10. The BERA repeatedly states that laboratory test conditions fail to adequately represent the conditions present at the Site (especially in terms of UV light). While this does produce some uncertainty in the risk analysis, it does not necessarily mean that lines of evidence based on laboratory testing should be given "very low weight" in the weight-of-evidence approach, especially since modeled exposures were used for several measurement endpoints.

Response

We believe that laboratory bioassay conditions do not adequately represent conditions at the Site [(and especially in terms of UV light which was the only bioassay results accorded "very low weight" (Section 6.2.1.4)], and therefore should be accorded less weight of evidence than the field studies.

Bioassays attempt in small bioassay beakers (300ml) to replicate a number of factors that an amphipod or other epifaunal species would experience in the field. Among the many field variables which cannot be adequately replicated in the laboratory are:

1) exposure to a composite of the top 6 inches of sediment PAH levels when most organisms in Site sediment are only exposed to the surface or top 2 – 4 cm of sediment where PAH concentrations are normally lower;
2) use of laboratory reared organisms that have not undergone natural selection or been adapted to conditions in the real world;
3) absence of the nepheloid layer which in the field moderates light exposure; and
4) limited space for normal behavior and avoidance of competition for resources.

11. The risk assessment itself is very long winded in its general overview of the fate and effects of various COPCs, but comparatively slim in its actual evaluation of the site data. As an example, an extensive sediment toxicity testing effort is simply summed up in a table of NOECs and LOECs. Pages of discussion of UV effects are dismissed with a simple Athat=s an uncertainty, not an effect.@ There is little attempt to integrate different information into a more thorough assessment. For example, the bioavailability information provided by the bioaccumulation testing is not considered as a tool to help interpret the results of the sediment toxicity tests. It is very surprising that nowhere in this document or the appendices is there a single graph showing the relationship between a toxicity parameter and sediment chemistry. We would have thought that would be the first thing to do, and much more robust than the (sometimes arbitrary) assignment of NOEC/LOEC values. Relying solely on NOEC/LOEC values from hypothesis testing rather than looking at exposure/response relationships seems like a step backwards in risk assessment methodology.

Response

We believe that the discussion in Section 5 of the BERA covers the most important caveats necessary for the public to understand the mechanisms of toxicity of PAHs, UV light, and photoactivated PAHs and is therefore relevant. The effects of the bioassays were summarized as NOECs and LOECs because:

1) The USEPA bioassay guidance (USEPA 2000) emphasizes hypothesis testing, not curve fitting;
2) The previous ERA by SEH expressed their data as NOECs and LOECs, not LC20s or LC50s, so we followed that precedent; and
3) Most of the risk management decision making we are aware of focuses on the NOEC and LOEC endpoints.

The bioaccumulation studies were not used as a tool to help interpret the results of the sediment toxicity tests because these results predict toxicity to L. variegatus where survival of H. azteca was high. See Table 1.

12. With regard to effects on the benthic community, the risk assessment concludes that there is evidence for risk, which we agree with. However, we do not agree with several aspects of the analysis that are used to estimate risk thresholds. If all that is needed from this document is a decision as to whether there is risk to benthos, then the details of the risk estimation methods may

not need to be decided now. However, if the analyses in this risk assessment are to be the basis for deriving cleanup goals, then the analysis needs revision and possibly supplementation, depending on the level of resolution needed to design remedial actions.

Response

Additional clarification has been presented in these responses which we believe will justify use of both the sediment toxicity testing results and benthic community assessment results for setting cleanup goals.

13. The sediment toxicity analysis is overly simplistic and arrives at risk thresholds well above concentrations shown to be dramatically toxic to benthic organisms. The same data could be used to justify a risk threshold roughly 10 times lower than what the authors have proposed. See Figure 1.

Response

NSPW does not agree with USEPA's conclusion that the proposed NOEC is greater than concentrations shown to be toxic. USEPA's Figure 1 does not appear to accurately portray all of the available bioassay data for H. azteca. In particular USEPA's Figure 1 does not appear to include any of the no effects data from the SEH bioassay studies and also draws conclusions that are biased by two URS data points from the 28-d bioassay that had high reference station mortalities (See further discussion below). Had USEPA included the SEH no effect data it would have been apparent, (as shown in Figure 1, which includes all the 28-d data from both SEH and URS as well as the 10-d URS H.azteca data) that a reasonable estimate of the NOEC, as reported by SEH, is 9,978 ug PAH/gOC. Based upon this, the RAO proposed by URS (5310 µg PAH/gOC) was very conservative.

The delineation of toxicity to H. azteca is complicated by two issues: 1) the range of PAH concentrations tested in the wood mulch was limited since there was no significant mortality at any wood stations, and 2) there was high mortality in reference station sediments in the URS 28-d tests. As discussed below, we believe both of these datasets should be accorded low weight of evidence in making risk management decisions as there are data from other bioassay tests with less uncertainty.

The data from both SEH and URS wood stations can not be used to estimate risks to the benthic community because other than the one treatment in the SEH tests there was no significant mortality at any wood station. Figure 2 shows the organic carbon-normalized SEH (2001) 28-d data in wood and results of the two tests conducted by URS in wood using H. azteca. In the wood tests, the highest concentration tested by SEH was 2,873 µg PAH/gOC and the highest concentration tested by URS was 250 µg PAH/gOC. No effects were found at the wood stations in either the URS or SEH bioassays except at the SEH 100% concentration. However, this result is suspect because the concentration, 1,582 ug PAH/gOC, is approximately ½ the TOC concentration

as that of the 50% dilution, the opposite trend of what is expected, and survival was 86.2 to 90% at higher PAH concentrations (1,963 µg PAH/gOC (10% wood) and 2,873 µg PAH/gOC (50% wood), respectively; Figure 2). USEPA also identified TOC measurements in the SEH sand dataset as an issue (Comment 31). If this one data point is considered an artifact of TOC measurement, there was no mortality at any wood station in multiple tests in different years. If this one data point is eliminated, there was no mortality at any wood station in multiple tests in different years.

Because of this the results for Wood bioassays are not a good basis for risk management decisions and the remainder of this discussion will focus on bioassays conducted at sand stations.

Although the concentrations tested at sand stations encompass the range of PAH concentrations at the Site, the data from the URS 28-d bioassays at sand stations should not be used to estimate risks to benthic communities because of high rates of mortality at the reference stations.

Between both the URS bioassay program and the SEH program (it should be noted that these bioassays were conducted by the same laboratory, LSRI, and supervised by the same personnel, Matt TenEyck), four 28-d bioassays and one 10-d bioassay were conducted with H. azteca using normal laboratory light. One 28-d bioassay was conducted by SEH in 2001 and three 28-d[2] and one 10-d bioassays were conducted by URS in 2005. The results of the 28 day bioassays are depicted in Figure 3. A NOEC of 9,978 µg PAH/gOC and a LOEC of 4,842 µg PAH/gOC were reported by SEH (SEH 2002). However as noted by USEPA the LOEC value is suspect because of what appears to be an error in the TOC measurement (See also Comment 31). When that 50% TOC concentration is estimated by linear regression between the 25% and 100% OC-normalized PAH bioassay data an estimate of the LOEC is 14,396 µg PAH/gOC (both values are shown in Figure 3). Once corrected and plotted with the other data a reasonable dose-response curve results for the SEH data. Figure 3 also shows that the two 28-d URS H. azteca tests resulted in NOECs substantially lower than the SEH NOEC. However, this outcome was associated with very low survival in the Reference Sand station sediments during all tests.[3] This is evidence that factors other than sediment PAH concentrations may have affected the outcome of the URS 28-d tests. Based upon this result the URS 28-d bioassay data are not a reliable basis for risk management decisions.

---

[2] A third URS bioassay was stopped after 21-d due to reference station mortality. This data is not plotted, but is mentioned because, like other tests in 2005, reference station survival was below acceptance criteria (80%) while the silica sand and SQT5 were acceptable.

[3] It should be noted that of the Site Sand stations three of the four Site Sand stations had significant mortality in the first 28-day test and two of the Site Sand stations had significant mortality in the second 28-d test. While two stations in both tests, SQT1 and SQT7 had elevated levels of carbon normalized PAHs, a third Site Sand station, SQT3 also had significant mortality in the first 28-d test. SQT3 had only 42.7 µgPAH/g OC and also had approximately 40% TOC. Total PAHs on a non-normalized basis at this station was 17 µg/g. While there was no significant mortality at Sand Station SQT3 in the second 28-d test, there is insufficient consistent information to support a conclusion of whether the same variable that apparently affected the reference stations did or did not affect some or all of the Site Sand stations.

However, if the URS 10-d data is combined with the SEH 28-d sand data (corrected as described above) there is good agreement between the two tests (Figure 4). Although these tests are of different durations it appears that the organisms received all or most of their dose during the first 10-d of exposure and there was little further toxicity. In both of these tests reference station survival was also good. As part of the USEPA comments to the BERA Erickson calculated an LC50 of 12,748 µg PAH/gOC and an LC20 of 9,593 µg PAH/gOC using just the URS 10-d data. These results are also consistent with the LC50 of 10,500 µg PAH/gOC reported by Driscoll et al. (1997) for H. azteca exposed for 10-d to fluoranthene. Since the toxicity of all PAHs occur through the same mechanism and is additive, the LC50 of mixtures should not be different from that of fluoranthene when normalized to OC. Based upon these results NSPW believes the proposed RAO is conservative.

USEPA commented that the data provided by URS can be used to justify a ten-fold lower RAO. It is not clear from the comment how this 10-fold lower value was derived, but several factors are apparently involved: 1) as discussed above, USEPA did not consider the SEH no effect data which shifts the dose-response significantly to the right (Figure1), and 2) USEPA seems (see Comments 20, 43 and 44) to have relied upon the equilibrium partitioning sediment benchmarks (ESB) document (USEPA 2003) that we believe is overly conservative for freshwater organisms such as H. azteca, particularly when the sediments being considered have TOC in various forms including in soot and coal, that result in adsorbed contaminants being less bioavailable than other sediments.

The ESB benchmark is based on toxicity data for the marine amphipod Rhepoxinius abronius, not H. azteca. Figure 5-4 of the ESB document indicates that R. abronius is more than 4-fold more sensitive than H. azteca, and Lee et al. (2001) reported that effects concentrations predicted from R. abronius toxicity data did not produce toxicity in H. azteca. The ESB toxicity benchmark is also based on a critical body residue (CBR) for R. abronius of 15.8 µmol/g lipid, which is about one-half the CBR of 35.3 µmol/g lipid used in the target lipid model of DiToro and McGrath (2000). In turn, the target lipid model CBR is consistent with the CBR data for other freshwater organisms, including H. azteca. Lee et al. (2002) reported that the CBR of fluoranthene for H. azteca ranges from 37 to 311 µmol/g lipid and Driscoll and Landrum (1997) reported even higher body residues of fluoranthene, ranging from 66.3 to 66.9 µmol/g lipid, in H. azteca that survived for 30-d of exposure at the LC50 concentration. In addition, Lee at al. (2002) also reported fluoranthene CBRs of 39 to 134 and 40 to 190 µmol/g lipid for the freshwater amphipod Diporea spp. and the fathead minnow, respectively. Therefore, the ESB model is inappropriate for H. azteca and freshwater organisms, in general and particularly when there are forms of TOC that render PAHs associated with the sediment less bioavailable.

It would also appear from Comments 20 and 42, that USEPA believes that the proposed RAO should be divided by an uncertainty factor of 2 in order to account for the difference between the 23 PAHs measured by URS and the 34 PAHs measured in some USEPA EMAP monitoring programs. However, as discussed in Response 42, this is, likewise, inappropriate.

14. The authors elected to eliminate UV-induced PAH toxicity from the effects analysis. There is little meaningful justification for this. It should be included as part of the analysis; the authors concerns about the applicability of this information can be addressed in the uncertainty discussion. At present, not only has UV-induced PAH toxicity been removed from the effects discussion, it is not even discussed as an uncertainty.

Response

We will present the UV light toxicity data as recommended.

15. The accumulation data for *Lumbriculus* are not used to their full effectiveness, as a means to explore bioavailability issues that could underlie all of the benthic community assessment. Some of these data contradict assertions made in the body of the assessment.

Response

Since no specific examples or text references are provided in the comment, we cannot respond to this comment.

16. There is a heavy emphasis on the benthic community study as being the Astrongest@ line of evidence and not providing clear evidence of in situ effects on the benthic community. While the conceptual rationale for this is reasonable, it assumes the study has the discriminatory power to detect differences. The degree of variability observed, both within and between sampling locations, brings this very much into question. If the benthic community study has low power, then it is prone to underestimating effects and is in fact a weak line of evidence rather than a strong one.

Response

There is indeed great variation among sampling locations, and this is why we removed ("partialed" out) the variation due to sampling location which included substrate effects in our Analysis of Covariance (ANCOVAs) before looking for effects of PAH. Our null hypothesis was that including carbon normalized PAHs (NOCPAH) and total PAHs (TPAH) in the model produced no increase in R^2 (variation explained) after removing the variation explained by among station variability. Cohen (1988) posits increase in R^2 of 0.02, 0.13, and 0.26 as 'small', 'medium', and 'large' effect sizes, respectively. Using these effect sizes in Cohen's own program, SamplePower 2.0®, with 1, 55 degrees of freedom (df) our analysis reveals an a priori power to reject the null hypothesis when the alternative hypothesis of a 'medium' effect is true of 0.819. The power to detect a 'large' effect is 0.993 or almost certainty. So based upon these results if PAHs are having any effect on community structure, it is small. Power to detect a 'small' effect was 0.182,

This quite respectable power of over 0.8 to detect a 'medium' effect refutes the argument that the " ...study... is ... a weak line of evidence rather than a strong one." Our study and analyses did not have low power to detect effects of PAH on benthic community structure.

To further support our contention, we performed an a posteriori power analysis on the result of our ANCOVAS. While such analyses have limited usefulness, they can shed light on the frequency with which high variability makes detection of real effects difficult, using the actual partitioning of variability into that explained by PAH concentration (after removing among station variability) and unexplained variability. Table 2 presents the result of that analysis: power to detect a 'small' effect after partialing out station variability varied from a low of 0.34 to as high as 0.94. Average power was 0.57, or better than one chance in two of rejecting a false null hypothesis. Power to detect a 'medium' effect was 1.0: certainty for all 22 biological metrics! Again this indicates that the power of these statistical analyses to detect a change due to PAH was substantial.

What effect sizes were actually observed in this study? The third and last column in Table 2 presents these for the effect of both NOCPAH and TPAH, respectively. Average increase in R^2 for including NOCPAH in the model was 0.005; that for including TPAH in the model was 0.0036. The actual effect sizes observed were in general much smaller than Cohen's standard for a 'small' effect. Again, we argue that if PAHs are affecting community structure, the effect is small. Since we had good power to detect a 'medium' effect, failure to reject the null hypothesis in most of the ANCOVAs was not due to low power.

One may also argue that analyzing 22 different metrics of community structure constitutes a form of meta-analysis. Our repeated inability to detect an effect spread over a number of different measures of community structure further suggests the effect of PAHs on benthic community structure, if any, has to be small.

Sediment Toxicity Testing

17. The range of PAH concentrations in the authors studies did not, unfortunately, succeed in providing a good range of contamination near the effect threshold. There is a better than 20-fold gap in PAH concentration between SQT7 and SQT3 which bracket the purported threshold.

Response

Comment noted

18. The authors use a result from the 10-d sediment dilution study to establish the NOEC for *Hyalella*. This seems inappropriate given that it is mixing 10-d and 28-d results, and other samples in the 10-d dilution study with PAH concentrations higher than SQT7 do not show effects, indicating that the sensitivity of the 28-d test is greater than the 10-d test, as might be expected. It also ignores an effect concentration of around 1500 ug/g OC found by SEH (2002). Because of the

large gap in concentrations, we would assert that the threshold for effects is highly uncertain based on the available site-specific data. The text should be revised to reflect this uncertainty. Because sediment toxicity appears to be one of the main lines of evidence establishing the existence of risk, resolving this uncertainty may be a high priority for the FS.

Response

As discussed in Response 13, above, when the 10-d URS and 28-day SEH (after correction) results are compared, the dose-response slopes are nearly identical (Figure 4). This suggests that mortality to this laboratory organism occurs primarily in the first ten days. Based upon this and the Probit analysis conducted by Erickson discussed in Response 13, above, we believed it was appropriate to use both tests as the basis for developing H. azteca endpoints.

In fact, treatments in the dilution study with PAH concentrations higher than SQT7 do show effects. The 50% dilution of SQT1, estimated at 83.5 µg/g and 18,145 µgPAH/g OC only had 15% survival in the dilution study.

As discussed in Response 13 the 1582 µg PAH/gOC effect concentration reported by SEH for wood stations comes from the 100% sediment concentration in a dilution series in which the 50% dilution had almost twice the TOC of the 100% sample. USEPA also pointed out this problem in their comments to this BERA. We do not believe, therefore that these data should be relied upon.

The text will be revised to better reflect uncertainties associated with relying upon two different duration bioassay tests.

19. The authors chose to ignore data from "wood" stations in their derivation of thresholds for sediment toxicity. The stated reason for this is the belief that the wood matrix would be a poor sorbent for PAHs and, as a result, the organic-carbon normalized effect concentrations would be inappropriately low. While this is conceptually reasonable, it is not at all supported by the data. Bioaccumulation testing with *Lumbriculus* indicates that the bioavailability of PAHs in "wood" sites is very comparable to that in "sand" sites (see Table 1, attached). Accordingly, there is no evidence that PAH bioavailability misrepresented by OC-normalized PAH concentrations at wood sites. This is also important because it moves the LOEC down to circa 1500 ug/g OC with the inclusion of the Wood #1 site from the SEH (2002) study. This suggests three-fold greater risk than is indicated by the current analysis.

Response

As pointed out by USEPA, the bioaccumulation data include only one sand station. Although there is insufficient data to be certain about trends, when the BSAFs in USEPA's Table 1 are plotted the bioavailability of PAHs from one sand station appears to be lower than for three of the five wood stations (Figure 5). (It should be noted that SQT3, while originally labeled as a "Sand" station is considered a "wood" station because of the large quantity of wood in the sample. It is plotted as such in Figure 5). This suggests that the wood matrix may not bind PAHs in the manner predicted by equilibrium partitioning. Obviously more data would be needed to confirm this suggestion but

we do not believe it is illogical to expect PAHs associated with relatively large wood debris (up to ½" size) to be more bioavailable just based upon relative surface binding capacity.

Responses 13 and 18, above, provide the rationale for disregarding the proposed LOEC of 1500 ug PAH/gOC.

It is also important to note that while the bioaccumulation data are derived from L. variegatus, the toxicity data are derived from tests with H. azteca. Table 1 shows the hazard quotients based on bioaccumulation and those derived from the bioassays. As discussed in Response 11, survival was very high in the 28-d H. azteca bioassays where the bioaccumulation results predicted toxicity, e.g. in SQT5. Based upon these results we believe use of bioaccumulation results as the basis for the RAO is overly conservative.

In any event NSPW will not be proposing a "wood" sediment cleanup goal because the high variability in distribution, abundance and characteristics of the "wood mulch" make it impractical.

20. The authors proposed LOEC/NOEC values do not seem well founded when one looks at the totality of the site sediment toxicity testing done by either SEH or URS. This is shown in Figure 1. What one sees is a) there is evidence for toxicity at PAH concentrations considerably below the LOEC/NOEC proposed by the authors; b) there is a large range of PAH concentrations not represented by the data in hand; and c) there is nothing in the site data to suggest that the risk benchmark proposed by EPA using equilibrium partitioning is not applicable to the site. This USEPA value is 5 to 10 times lower than the concentrations suggested by the authors (the range is created by assumptions regarding the role of unmeasured PAHs; lower bound of USEPA estimate assumes that priority pollutant PAHs comprise about half of the total PAH mixture, based on other work at coal tar sites [see publications by Kreitinger et al.]).

Response

As discussed in Response 13, USEPA Figure 1 in the BERA comments appears to ignore much of the SEH 28-d data that had acceptable survival. When those data are added to USEPA Figure 1 the results support the NOECs and LOECs proposed by NSPW (Figure 6). If the results of the 10-d test are also included (Figure 4) it further supports this recommendation.

As discussed in Response 13 the USEPA sediment risk benchmark is 5 to 10 times lower than the proposed URS NOEC because it is based on a marine amphipod species that appears to be twice as sensitive as H. azteca on a body residue basis and 4.3 times more sensitive on a sediment NOEC basis.

Because effects levels for bioassays as well for the benthic community analysis were based upon the same 24 PAHs as were measured in the bulk sediment analysis, the 24 PAHs can represent all PAHs, measured and unmeasured. Only the assumption that the relative proportion of non-EMAP

PAHs to the total PAH list remains relative constant need be made. Furthermore, evaluation of the total PAH data for Site stations based upon the results for all SQT stations (Newfields 2006) shows that the difference between the 24 PAHs measured in organisms and bioassay sediments and the USEPA EMAP list of 34 PAHs is a factor of 1.09, not 2.0.as suggested by USEPA (see Table 3).

In summary the work of Kreitinger et al. does not appear to be relevant for this Site and the proposed RAO should not be divided by two as suggested by EPA in Comment 42, below.

21. Although it is not justified by the data, if the risk assessment discounts wood stations in evaluating risk, how would clean-up goals for areas with wood substrate be determined? Further, why would one include wood stations in the community analysis if one can't relate chemical concentration to expected risk?

Response

NSPW does not believe it practical to set clean-up goals for sediment containing high levels of wood mulch.

Wood stations were included in the analysis for the following reasons:

1) Because of the large quantities of wood mulch in the majority of Site sediments, there was a potential that it could be a significant determinant (or independent variable) of benthic community composition;
2) Analysis of covariance (ANCOVA) allows removal of the effects of potentially confounding variables, like wood, so that the contributions to variance in community structure due to wood will not confound differences, or lack of differences, between contaminated and reference stations;
3) The presence of wood mulch in Site sediments could also have influenced the outcome of sediment toxicity tests so it was necessary to see if there were differences in bioassay results due to the presence of wood before concluding that the results of non-wood Site stations were likely due to the presence of contaminants.

22. How can SQT3 be a sand station and have 40% organic carbon? Was the composition of the organic carbon in this sample verified? This is the OC range observed in wood stations; it does not sound like sand. This may be particularly important since SQT3 seems to be on the borderline of toxicity, even though the apparent OC-normalized PAH concentration is not that high. If this station had an OC concentration more typical of a sand station, then this might make sense.

Response

Station SQT3 was designated to be a Sand Station and after collecting some trial samples during the reconnaissance it was sampled as a sand sample. After the lab results were received, it was clear that actual samples had higher levels of organic carbon than was planned. The five bulk sediment replicates ranged from 17 to over 40% TOC.

The problems with the wood bioassays were discussed in Response 13 above. Because of problems in the SEH wood bioassay and the absence of any effects at wood stations in the URS bioassay, NSPW does not believe that wood bioassays are a good basis for judging the toxicity of these sediments. Furthermore, with regard to SQT3 being the borderline for toxicity, Stations SQT5 and SQT8 have higher OC-normalized PAH concentrations but showed no toxicity, possibly because OC-normalization does not appear to be appropriate for these tests.

23. Table 5-3 proposes a "no effect" level of 7257 ug/g OC for *Hyalella*. This needs to be clarified. The authors own tests show that a sediment with 6090 ug/g OC caused between 83 and 94% mortality. In addition, SEH found a sediment with 1580 ug/g to cause 97% mortality. This is in no way a "no effect" concentration.

Response

The proposed LOEC of 7,257 μg PAH/gOC was based on the 28-d tests, but the high mortality at reference stations prevented establishment of a 28-d NOEC based on the URS tests. Therefore, we used the 10-d NOEC data. However, since the 10-d NOEC (9,072 μg PAH/gOC) was higher than the 28-d LOEC, and these were 10-d, not 28-d tests, we selected the next lower NOEC concentration for the purpose of calculating an average NOEC. This value, 4,536 μg PAH/gOC, was averaged with the SEH 28-d NOEC(9,978 μg PAH/gOC) to produce the value of 7,257 μg PAH/gOC shown in Table. 5-3.

However, through this comment-response process we have now shown that there doesn't appear to be a difference between the SEH 28-d data and the URS 10-d data. As a result we now believe that the NOEC for H. azteca should be based on the average of 9,072 and 9,978 μg PAH/gOC, or 9,525 μg PAH/gOC. This NOEC is very close to the 10-d LC20 of 9,593 ug PAH/gOC calculated by USEPA by Erickson in the BERA comments.

As discussed in Response 13, the high mortality at a concentration of 6,084 ug/gOC reported by URS comes from the 28-d tests in which reference station mortality was also high. The effect concentration reported by SEH as 1582 ug/gOC is likely an artifact of the TOC measurement. We believe neither of these data should not be used as the basis for risk management decisions.

24. The failure of the authors to produce additional data on the toxicity of sediments to *Chironomus* leaves the assessment to rely only on previous data. In discussing these data (page 5-

18) the authors imply some uncertainty about effects at 3900 ug/g OC. However, it's worth pointing out that there was also 100% mortality at 4800 ug/g OC B so if anything, it's surprising that there wasn=t more mortality at 3900 ug/g OC, which is the opposite of questioning the existence of effects at 3900 ug/g OC.

Response

The reason we emphasized that the survival of Chironomus at 3,996 ug/g/OC was 90% of the control (73% survival at 3996 µg/g OC and 82.5 % survival in the silica sad control) was that even though it was significantly different from the control survival the survival was sufficiently high that it would have met the test acceptability criterion for a control sample, i.e. 70%. This essentially means that it was within the variability expected in these bioassays even in the absence of contaminant effects. Therefore, 3,996 ug/g/OC should be the NOEC, not the LOEC.

As discussed in Response 13, the PAH concentration associated with 100% mortality in the SEH bioassay tests is probably much higher than the 4,800 ug/g/OC identified by USEPA. As pointed out by USEPA in Comment 31, the value of 4,820 µg/gOC at the 50% dilution is probably an artifact due to a TOC measurement error. Our estimate of the LOEC based upon interpolation is much higher, 14,396 ug/gOC (see Response 13). Based upon this, the corrected NOEC and LOEC for Chironomus should be 3,996, and 14,375 ug/g/OC, respectively.

25. There are several references to spurious effects in the sand reference stations (e.g., 5-17), implying that site sediments may be affected by regional background contamination in addition to site-related contamination. However, it's worth noting that there was also indications of poor performance in some laboratory control treatments (formulated sediments) and, more significantly, there was no indication of spurious toxicity in sandy site sediments with lower PAH contamination. From this, it is not at all clear whether the results found in sand reference sediments have import to the assessment of site-related risks.

Response

There was less than acceptable survival of H. azteca in the formulated sediment performance control only during the second 28-d URS bioassay. In the first URS 28-d test, survival in the formulated sediment was 80%. The formulated sediments are an artificial substrate and the LSRI laboratory manager has reported indications of bacterial overgrowth and sulfide production in bioassays conducted with the formulated sediment. Kemble et al. (1999) also reported problems with formulated sediments using the same test protocols used by LSRI. However, the formulated sediment is only one of three performance controls and survival was always high in silica sand and West Bearskin Lake sediments. Silica sand is the standard control at the LSRI and West Bearskin Lake sediment is the preferred reference sediment of USEPA-Duluth. Since survival was always well above acceptance criteria in these controls, this shows that the high mortality at sand reference stations was not due to laboratory error as implied by this comment.

Mortality in sediment toxicity tests at four different reference stations, two of which were tested twice, ranged from 23.7 to 62.5% and the PAH concentrations associated with these adverse effects ranged from 1.43 to 152 ug PAH/gOC (14.3 to 1520 ug/kg @1%OC). This is below even the Threshold Effects Concentration (TEC) of 1610 µg/kg@1%OC. Based upon this result it is likely that these effects were due to something other than PAHs in the reference sediments.

In addition as discussed in Response 13 three of the four Site Sand stations had significant mortality in the first 28-day test and two of the Site Sand stations had significant mortality in the second 28-d test. While two stations in both tests, SQT1 and SQT7 had elevated levels of carbon normalized PAHs, a third Site Sand station, SQT3 also had significant mortality in the first 28-d test. SQT3 had only 42.7 µgPAH/g OC and also had approximately 40% TOC. Total PAHs on a non-normalized basis at this station was 17 µg/g. While there was no significant mortality at Sand Station SQT3 in the second 28-d test, there is insufficient consistent information to support a conclusion of whether the same variable that apparently affected the reference stations did or did not affect some or all of the Site Sand stations.

26. On Page 6-4, the conclusions regarding risk levels for benthos are inappropriate. First, as shown in Figure 1, the interpretation of the *Hyalella azteca* data are such that the proposed NOECs would be higher than concentrations that caused near complete mortality. The suggestion that the *Chironomus* test procedure is inadequate for assessing effects is concerning, as it has been evaluated through round robin testing and widely applied. I have no idea why the authors had difficulty conducting the test, but that does not mean that the species is irrelevant to the risk assessment. More importantly, this decision ignores data presented by SEH that clearly shows adverse effects at concentrations below the suggested NOEC values. Finally, the analyses by Ingersoll et al. on Great Lakes sediments used chronic *Hyalella* data but only 10-d exposures for midge, which may explain part of the trend they observed.

Response

We do not believe the conclusions regarding risk levels for benthos are inappropriate because, as discussed above, the higher mortality to H. azteca in the 28-d bioassays to which the reviewer refers can be attributed to unknown sources of toxicity that were documented at four different sand reference stations.

In addition as discussed in Response 13, above, Figure 1 in the USEPA comments to the BERA upon which this statement is at least partially based:

1) does not include all of the SEH 28-d no effect data; and

2) does not correct for apparent errors in TOC measurements (see USEPA Comment 31).

In addition the benthic community analyses showed:

1) Gammarid amphipods and isopods (both crustaceans that are phylogenetically and physiologically similar to H. azteca) were more abundant at Site sand stations than at reference sand stations suggesting that there was little, in any effect to these species;

2) Station SQT7 that was acutely toxic to H. azteca in bioassays had more gammarids than any of the four reference stations; and

3) The proportion of crustaceans + molluscs was higher at Site than reference stations.

These observations are consistent with the results of the bioassays that indicated that other variables other than PAH level in the sediment were responsible for the high mortality at Sand reference stations.

While the 20-d Chironomous survival and growth test has been fairly widely applied, the outcome of these tests is far from consistent. As reported in a recent study of round robin tests (Norberg-King et al. 2006) only 63% of the laboratories met test acceptability criteria for the Chironomus 20-d survival test, the same as that conducted by URS. The failure of the URS chironomid tests was explained in the BERA as the failure of the Chironomus larvae to settle on the substrate and their subsequent washing out of the bioassay beakers. This failure to settle is considered to be an effect of the very early life stage used by URS (the SEH tests used older life stages of Chironomus and for only 10-d).

At least two plausible explanations for this failure to settle were found in the primary literature: 1) De Haas et al. (2006) reported that larvae of C. riparius show a clear preference for sediments with high quality food and that food quality overruled sediments with higher toxicant concentrations. They also reported that the first-instar larvae that were used by URS, were more sensitive to contaminants than the older organisms used by SEH; 2) Edgar and Meadows (1969) reported that smaller C. riparius move more than larger C. riparius, that larvae could not build cases on sand substrates, and when provided only with sand substrates exhibited "a violent wriggling and swimming movement". This behavior and food selection preferences could allow them to be swept away during water renewal.

The benthic community data also call into question the results of the Chironomus bioassays. In contrast to the SEH-reported NOEC of 3,996 ug PAH/gOC for Chironomus, the benthic community studies conducted by URS in support of this BERA indicated that station SQT7 chironomid abundance equaled or exceeded that at three of the four reference stations. The PAH concentration at station SQT7 was 6,084 ug PAH/gOC.

Like Ingersoll et al. (1996), we also relied upon a 28-d test with H. azteca and a 10-d test with Chironomus, both conducted by SEH. We would therefore expect these results to be consistent with the trend he reported, i.e., that H. azteca would be more sensitive than Chironomus.

Photoactivated Toxicity

27. The authors have chosen to treat UV/PAH effects as an "uncertainty" rather than a component of the risk assessment. While we don't agree with this decision, the authors do not

include this issue in the uncertainty analysis, particularly with the notation that this omission errs entirely on the side of underestimating effects.

Response

The UV light results, as well as those where refugia from UV-light was provided, will be discussed in more detail in the final BERA.

28. One of the reasons given for discounting photo-activated PAH toxicity is that the UVB levels in the laboratory exposures were higher than those measured in the field. We are not sure from this comment whether the concern is that UVB levels were high enough to cause direct phototoxicity, or the excess UVB would contribute substantially to the photo-activation of PAHs. In either event, it seems a little odd for the authors to dismiss the studies as inappropriate when they were themselves responsible for designing and executing it. While we would accept that excess UVB would create an uncertainty in applying the study results, it is not at all clear that it is a legitimate basis for disregarding the entire mechanism.

It was also a little surprising that the lab light source would have that high a UVB intensity after filtering through glass, although UVB removal by glass is dependent on the thickness of that glass (not specified) and to some degree on its composition. In our laboratory exposure system, which uses the same UVA-340 bulb, we made measurements using the same model International Light 1700 meter used in the Ashland study. Inserting a piece of 1/8" frosted window glass reduced UVB to only 6.6% of its original intensity, while UVA was only reduced to about 30%. Put differently, the ratio of UVA to UVB without glass filtration was 3500:292 or about 8:1 (which is similar, but slightly higher than the roughly 10:1 ratio in sunlight). After passing a 1/8" sheet of window glass, this ratio 1066:19.3 or about 55:1, which is much higher than the 12:1 ratio measured during the laboratory UV assays. While the thickness of the glass used as covers is not described, it is very likely that using thicker glass filters could bring the ratio closer into line with nominal.

Response

Certainly either excess toxicity due to UVB alone or excess photoactivation were possible outcomes of our inability to reduce UVB to the modeled objectives. It should also be noted that NSPW considered (and still consider) this line of evidence to be research in the preliminary stage and that it introduces unnecessary uncertainty and conservatism into risk management decisions for the Site.

Nevertheless, we did not dismiss UV light as a possible mechanism of increasing the toxicity of PAHs. We merely presented the commonly stated opinion of USEPA (Swartz et al. 1997; Bell et al. 2004; Diamond et al. draft) and other researchers (McDonald and Chapman 2002; Hatch and Burton 1999) that UV light is unlikely to cause the same effects in the field that are observed in the laboratory. Even USEPA guidance on evaluating the potential effects of photoactivation of PAHs

(USEPA 2003) states, "If photoactivation of PAHs is ecologically relevant, it is probably most significant primarily for organisms that inhabit very shallow or very clear water."

It should also be noted that LSRI used the same UV-light system that was used for the SEH tests. The glass used to filter the UVB light was 3 mm thick, nearly identical to the 1/8 inch used by EPA, but it was not frosted. The UVA:UVB ratio was about 13:1.

29. While it ends up having little impact on the final UV levels used in the UV studies, the analysis used to derive the UV exposures contains errors. Most significantly, the authors used an average of readings taken at 1000 and 1400 hours to create an estimate for 1200 hours. Because solar noon is at approximately 1300 hours at Ashland under Daylight Savings Time, and because solar irradiance is for all practical purposes symmetrical around solar noon, the 1400 hour readings would have been the appropriate estimate of the 1200 hour readings rather than the mean of 1000 and 1400. The estimated clear sky irradiance curve was re-modeled after correcting this error, and obtained Figure 2. While this curve lies above the curve used by the authors, a more detailed algorithm was used to estimate total daily UV dose, by using the same basic approach as the authors, but using 0.1 hour time steps. This more detailed averaging altered the total UV exposure estimate by the nearly exact opposite amount as recalculating the daily irradiance curve. The net effect is that recalculated values were very close to the target values used originally derived. It is brought up here only for the record in case similar calculations are used later in the RI/FS process.

Response

In developing the light regime for the UV tests there was considerable interaction with USEPA's Duluth lab. We assumed that this estimate of 1200 hours light intensity had been critically reviewed [See Appendix 1 to LSRI bioassay report (In Attachment 2 to Appendix B of the BERA)]. We will follow this suggestion if future work is necessary.

30. The depth for which UV exposure was estimated is listed in the LSRI report as being for SQT2, the deeper of two stations evaluated, but we believe it was actually for 232 cm depth, which was that for SQT1, the shallower station. As explained at length in the background material, incident UV is a function of depth, and the potential for photoactivated toxicity varies with depth accordingly. This can be incorporated into the analysis using a reciprocity assumption, discussed in the background material and well supported by experimental work. This relationship says in essence, that if you double the UV exposure, the PAH body burden associated with toxicity will be halved. Alternatively, if you halve the UV exposure, the PAH body burden required for equivalent effect will be doubled.

Potency of the Site PAH mixture can be estimated by comparing the responses in the dilution studies of SQT1, which was tested under lab light, under UV light, and under UV light with leaf plugs added as a potential source of shading. The LC50 and LC20 concentrations can be estimated independently for each, as shown Table 2. See Attachment 1 for further details on this analysis.

As narcosis and photo-activated toxicity occur by different mechanisms, the sediment toxicity under lab light can be reasonably assumed to be independent from the photo-activated toxicity. As such, one can estimate the effect of photo-activated toxicity alone by removing the binomial probability of mortality under lab light from that under UV light (with or without shading). These estimated values are also shown in Table 2. This analysis indicates that the presence of leaf plugs reduced the effective UV exposure by about 40%, but UV exposure was not eliminated.

From these corrected EC20 values, one can estimate the sediment PAH concentration that would be associated with a 20 percent effect at depths other than the one simulated by the lab assays (SQT1). These values are shown in Table 3. While little additional toxicity from photo-activation would be expected at stations deeper than SQT1, for areas of the site shallower than SQT1, risk thresholds for PAH contamination could be expected to decrease dramatically. Development of remedial goals for areas substantially shallower than 232 cm may need to consider more rigorously the influence of photoactivated toxicity in establishing risk to benthos.

Response

The reviewer is correct. The UV-light measurements used in the bioassays were conservatively based on the shallower of the two bottom depths where UV-light was measured. The full method, correctly stated, is found in Appendix 1 of to Attachment 2 of Appendix B. The URS UV-light exposure regime was based on site measurements that were adjusted to approximate summer solstice light conditions at 75% light incidence. This regime was arrived at following discussions with USEPA-ERL Duluth.

Figure 7 shows that there was a similar increase in toxicity in both SEH and URS studies due to UV light and no substantial difference in the NOEC and LOEC between the URS 10-d UV tests and the SEH 28-d UV tests. This is similar to the results of the 10-d URS and 28-d SEH tests in normal light discussed in Response 13, where test duration had no apparent effect. However, in order to better mimic field conditions, URS also conducted 10-d tests where leaf plugs were provided so that H. azteca could avoid the UV light. As shown in Figure 8, when the 10-d URS UV + Refugia results are compared to the combined URS and SEH normal light dataset, there is little difference between the two curves. In the USEPA BERA comments, Erickson 2006 estimated the difference between the UV Refugia results and the combined URS 10-d and SEH 28-d normal light dataset at the LC20 to be about 13% (See USEPA's Table 2: 9592 vs. 8359). Given that there are numerous differences between the PAH concentrations, organisms, habitat, and behavior between the laboratory and field, we do not think this small increase in toxicity would be measurable in the field and even less likely that an effect this small would be manifested at the population or community level.

We agree that development of remedial goals in depths less than 232 cm should consider that the effects of UV light are greater than were measured in the laboratory bioassays.

31. Somewhere in the document or its appendices it was suggested that there was inconsistency in response between the UV studies reported by SEH and those conducted by the authors. The data was reanalyzed and it was found not to be true. The PAH concentration data for the SEH data are confused by what appear to be spurious OC measurements (OC contents are not monotonic across the sediment dilution series). Problems with sediment OC measurements are not uncommon, particularly when the analytical instrument uses small sample sizes and the sediments contain large organic particles. However, if one simply expresses exposure as percent sediment, and then models the LC50 measured in the presence and absence of UV light, one gets a ratio of 3.4 for the sand series and 3.7 for the wood series. After correcting for the slightly higher UV in the SEH studies, one would expect a ratio of 2.6 based on the URS dilution studies. We would argue this is a pretty good agreement, rather than inconsistency.

Response

This was Section 2.1.1 of Appendix B Attachment 2 Volume II. in the BERA . We used the data provided in the SEH report. Our comment was that a spurious TOC measurement led to the illogical conclusion that the LOEC was lower than the NOEC on an NOC basis. If the SEH report contained "spurious" TOC data, this point was not made in the report or in agency reviews and the effects concentrations were not corrected.

However, as discussed in Response 13, correcting for spurious TOC of the 50% dilution as discussed above, the LOEC concentration should be 14,396 ug PAH/g-OC.

32. The LSRI report also states that the modeled UV assumption was that the sky was 75% clear. While the 75% correction is accurate, the rationale is slightly different. The assumption is that the effect of cloud cover and other weather would, over time; result in an average incident UV that is approximately 75% of the clear sky value. As such, this estimate is an average of expected exposure, not a worst case. The worst case (10 consecutive sunny days) would be expected to be roughly 1/3 more potent than the condition tested.

Response

We agree.

33. Section 5.1.2.1 lists (bottom of page 5-10) UVB as the most damaging UV radiation. This is true for some toxic mechanisms, but not all. This distinction is further modified by the specific environment, such as for the example of water column attenuation discussed here. The mix of mechanisms discussed with respect to UV create problems with generalizations like this.

Response

In an effort to simplify the discussion for all readers the discussion of potential effects of UVB was somewhat generalized in Section 5.1.2.1 where the effects of UV light alone were discussed. We

stated, "Of the wavelengths reaching the earth's surface, UVB is the most dangerous to aquatic life." Were this not a true statement, life on earth would be very different. We believe this level of discussion is necessary in order not to gloss-over the very substantial separate effects of UV relative to the effects of UV and PAHs.

The effect of water column attenuation is discussed on page 5-11. The effects of UV activation of PAHs are discussed beginning on page 5-13.

34. Page 5-15 B AIn field samples, it may be difficult to separate the effects of UV light and PAHs from those of the multiple environmental contaminants inevitably present.@ Why any more so than any other endpoint? In the studies conducted, the relationship of toxicity enhancement by UV and sediment PAHs seems pretty clear.

Response

We were merely pointing out that it may be difficult to distinguish the subtle effects of UV light, if any, from those due to a number of other environmental variables that potentially affect ecological receptors in the real world. We did not suggest this statement couldn't apply equally well for other lines of evidence. As an example, generic sediment quality benchmarks such as the "consensus-based sediment quality guidelines" that are based upon correlation of bulk sediment chemistry characteristics with observed impairment of ecological receptors or communities, and not on causation, do not separate out the effects of specific contaminants.

We disagree that "the relationship of toxicity enhancement by UV and sediment PAHs seems pretty clear" (See Figure 8). While there appeared to be enhancement of PAH toxicity in some of the bioassays, these results cannot be necessarily extrapolated to the field. In fact, the assessment of benthic community structure conducted for the BERA indicated there was no consistent significant differences between stations that could have been related to PAH level, much less the incremental effects of UV light on sediment PAHs.

As discussed in Response 30, when all of the available toxicity data for H. azteca is plotted, the normal laboratory light and UV-light data with refugia are very close (Figure 8). This shows that the effects of UV-light plus PAHs are not easily distinguished from those of PAHs alone, even in the laboratory.

35. Page 5-15 B ATherefore, the ability to create photosensitivity under laboratory conditions does not necessarily mean that toxicity will occur under natural conditions.@ This text seems to be downplaying the relevance of laboratory exposures to estimate effects in the field, but the mechanisms being discussed immediately previous to the statement (partitioning, burrowing into sediment) are operative in laboratory studies of the type conducted in this study, and are therefore not relevant to the issues at hand.

Response

We presume USEPA meant that the mechanisms discussed are relevant to the field. While it is true that these mechanisms occur both in the laboratory and in the field, the extent to which they occur differs significantly. In their natural habitat, mobile animals such as amphipods move more than the two inches allowed in the bioassay beakers and have access to more food and shelter. In addition, in the field, these organisms are exposed to only the surface of the weathered sediment, not the composited top 6 inches of sediment that, according to PAH analysis, as a whole contains largely un-weathered PAHs.

The results of the benthic community assessment conducted for the BERA confirm that the extrapolation of laboratory bioassays to the field is highly uncertain.

36. Bottom of page 5-15 B this discussion relies heavily on the McDonald and Chapman paper, the substance of which is far from a consensus viewpoint. More specifically, this discussion repeats arguments used by McDonald and Chapman which relied on quotes used without their original context in a way convenient to the arguments McDonald and Chapman were anxious to advance. For example, the Swartz study involved primarily deepwater sites, where one would not necessarily expect photoactivation because of the combined effects of water column attenuation and the burrowing habits of these particular amphipods. McDonald and Chapman present this as though photoactivated toxicity was expected but not observed. The quote from the Diamond paper does not actually appear in the published manuscript (Diamond et al. 2003). Although it was present in an earlier draft provided as a courtesy to Dr. Chapman, the quote as presented in their paper is also taken out of context; the adjoining text in the manuscript provided them went on to discuss the fact that exposure used in the experiments was not the same as that received in the field, but that other organisms with comparable sensitivity but higher UV exposure may be at risk. For the amphipods at that particular site, lower PAH exposure and high water column attenuation would be expected to protect these organisms; these factors have been accounted for in the study design used for the Ashland site, so the statements by Diamond et al. are not relevant to the situation at hand.

Response

Without getting into a debate about what these authors said or whether the authors took facts out of context, NSPW believes it is clear that there is substantial uncertainty in the scientific community regarding the effects of photoactivation of PAHs on ecological receptors in the field. This is an area of ongoing research and there are no generally accepted conclusions as to its importance. Even USEPA guidance on evaluating the potential effects of photoactivation of PAHs (USEPA 2003) states, "If photoactivation of PAHs is ecologically relevant, it is probably most significant primarily for organisms that inhabit very (emphasis added) shallow or very (emphasis added) clear water."

The URS 10-d bioassays conducted with refugia support Swartz et al. (1997), McDonald and Chapman (2002) and USEPA (2003) in that they show that the addition of UV light did not have a

substantial additional effect (about 13% at the LC20 according to USEPA's Table 2) over that caused by PAHs alone (Figure 4). We concur with USEPA (2003) that sediment benchmarks that do not account for UV-light "may" be under-protective, but believe this under-protectiveness is more than balanced by the uncertainty associated with extrapolating this small difference in the results of the laboratory bioassays using artificial UV light to the field.

37. Section 6.2.1.4 and 6.2.2.3 B Dismissal of a line of evidence should require something more than vague references to beaker size effects (not discussed anywhere else). This is particularly true when sediment testing included the provision of refugia (leaves). While the UVB may be an uncertainty, its seriousness is not defined and, further; it errs to the conservative side as the rest of the document touts as the way uncertainties were handled. This is not an appropriate justification for discarding this line of evidence.

Response

We did not dismiss this line of evidence We said "Although the bioassays conducted under UV light indicated effects thresholds at lower concentrations of PAHs, the small bioassay beakers (350 ml), inability to replicate a number of factors that an amphipod or other epifaunal species would experience in the field, and the fact that H. azteca were exposed to approximately eight times the UVB light measured in the field renders the results of the UV bioassays highly uncertain. Very low weight is placed on this line of evidence.". The beaker size is only one of many factors that call into question the relevance of extrapolating the results of UV laboratory tests to the field. See response 40 below.

38. Table 6-4 B the discussion of uncertainties does not include the potential underestimation of risk arising from discounting the potential for photoactivated toxicity.

Response

The table will be revised and expanded to address this uncertainty.

***Lumbriculus* bioaccumulation study**

39. The document asserts that wood chips present in some sites exaggerate bioavailability; the data from the accumulation study does not bear this out, as BSAFs at wood sites are comparable to those at sand sites.

Response

See Response 19, above.

40. The risk assessment asserts that using site-specific BSAFs is a "conservative" assumption. It would seem that these site specific BSAFs are in fact the best estimate of the true value, not a conservative assumption.

Response

The BSAFs obtained with L. variegatus are site-specific only in that they were obtained using Site sediments. The bioavailability of PAHs in the L. variegatus bioaccumulation test is at least partly an artifact of the USEPA test procedures. When samples are composited for bioaccumulation tests (and bioassays) from the top 6 inches of site sediment, higher concentrations from deeper sediment depths are mixed with lower concentrations from the surface layers where most organisms live. This is important because Harkey et al. (1995) also showed that "uptake rate is primarily driven by the first few sampling points [i.e., first few days] during a bioaccumulation assay". This means that, until the overlying surface waters leach the PAHs from the surface layers, bioassay and bioaccumulation organisms are exposed to much higher concentrations than they would be in the field.

Higher bioaccumulation also occurs during tests with L. variegatus because they are not being fed for 28-d and lose weight and lipid content. However, lipid content is not necessarily lost in the same proportion as total weight. The worms analyzed at the beginning of the bioassay contained an average of 2.44% lipid and this decreased to 0.5% at one station and less than 2% at most stations after 28-d. This alone causes an apparent increase in body residue although the relative concentration of PAHs in the lipid may not have changed.

Another possibility for the extraordinarily high bioaccumulation found in these worms is that the depuration period may have been insufficient for them to clear their gut contents. Harkey et al. (1995) report that only 2% of the body residue is released from worms during the 24-h depuration period. If the worms were narcotized, the active portion of depuration would presumably take even longer.

These factors could result in exposure conditions that most benthic invertebrates would not be expected to encounter. Furthermore, since the littoral zone at the site is covered by rip-rap, higher trophic levels would not be expected to consume worms. Most of the invertebrates consumed by wildlife are probably emergent chironomids or other aquatic insects that do not ingest sediment and leave a large portion of their body residue behind when they molt to aerial forms.

41. There is mention that soot carbon data were collected, but I did not see that in the documents. These data should be compared to the bioaccumulation data to see how they relate to observed bioavailability. Only one site, SQT8, shows accumulation consistent with of having bioavailability being affected by black carbon (total PAH BSAF = 0.15); all others have BSAF>1.

Response

Soot data appear in the Newfields forensic report (Newfields 2006) on Table 5. Coal, coke, slag and tar data are in Table 4 of the same report. See this referenced on page 5-4 of the BERA, "In addition to wood mulch another source of organic carbon in site sediments is soot. Soot made up as much as 10% of the sediment at some of the sediment stations (Table 5, Newfields 2005)."

Based upon these results SQT4 and SQT5 had 10% or greater soot and SQT8, 1.6%. SQT8 had 5.5% coke.

42. The PAH coverage in this study included only a couple alkylated PAHs. Other data (see papers by Kreitinger and co-workers) suggest that in other coal gasification waste sites, alkylated PAHs make up roughly half of the total PAH mass. To account for this, the NEBR should be divided by 2, or some other appropriate correction factor.

Response

In fact, of the 24 PAHs quantified in the RI, five were alkylated. The 24 PAHs on this list are the same as the ones NOAA uses for its Status and Trends program and several of the EMAP programs (See USEPA Guidance on PAH mixtures [USEPA 2002]). Other EMAP programs, including the Carolinian and the Louisianan monitor 34 PAHs. In the forensic study conducted by Newfields (2006) a total of 57 PAHs were quantified including 27 alkylated PAHs. The average ratio of the 34 EMAP PAHs to the 24 PAHs measured by was approximately 1.09 (Table 3).

In addition Newfields (2006) concluded that these Site samples were enriched in parent PAH compounds. This is typical of pyrogenic PAH sources. We could find only two abstracts with Kreitinger as the first author, but this and other papers by Hawthorne, Kreitinger et al. concluded that the USEPA equilibrium partitioning model overpredicted toxicity and that only the rapidly released phase PAHs (i.e., 2-3 ring PAHs) of MGP plant sediments contributed to toxicity (Hawthorne and Kreitinger 2003; Hawthorne et al.2005). The rapidly released fraction is obtained by supercritical carbon dioxide fractionation[4] for 40 to 120 min and Hawthorne et al. (2005) have shown that this extraction is comparable to that obtained with 120-d of extraction in water. Therefore, surface sediments that have been leached in water for decades at the Site should be relatively non-toxic if not mixed with deeper sediments where the rapidly released fraction has not been so leached. Hawthorne et al. (2006) also reported that, whereas the USEPA ESB model recommends multiplication of the 13 most commonly measured PAHs by 11.5 to estimate the ESB with 95% confidence, a factor of only 4.2 is required. There was no discussion comparing the 23 parent and alkylated PAHs measured by URS to the additional 11 alkylated PAHs measured in the EMAP program. However, Ozretich et al. (2000) reported that although modeled toxic units increase substantially when all these additional alkylated PAHs are included, they contributed little to the actual measured toxicity of the sediments at a former wood treatment facility.

Finally, as discussed in Response 20, because effects levels for bioassays as well for analysis of the benthic community data were based upon the same 24 PAHs as were measured in the bulk sediment analysis, the 24 PAHs as a can represent all PAHs, measured and unmeasured. Only the assumption that the relative proportion of non-NOAA PAHs to the total PAH list remains relative

[4] Supercritical carbon dioxide only extracts non-polar organic chemicals and does not affect other sediment components. Extraction by this method mainly removes 2-3 ring PAHs such as naphthalenes and leaves 5-6 ring PAHs unchanged.

constant need be made. Table 3 shows that the difference between the 24 measured PAHs the 34 EMAP PAHs is less than 10%.

43. USEPA's sediment ESB for PAH mixtures used only PAH toxicity data in an analysis similar to DiToro et al and arrived at an NEBR specifically for PAHs that is approximately 40% lower than DiToro et al. This is not acknowledged in the document.

Response

See Response 13, above.

44. Page 6-5 B The NEBR is exceeded beyond SQT1 and SQT7, particularly if you reduce the NEBR to account for unmeasured chemicals and/or use the USEPA ESB NEBR instead of DiToro et al.

Response

See Response 42, above.

45. Page 6-23 B Text regarding the bioaccumulation study suggests that the site specific bioaccumulation studies are somehow questionable because BSAFs greater than 1 were found. The DiToro estimates are not bounds on theoretical possibility, they are estimates based on specific assumptions regarding partition coefficients. If the partition coefficients at this site are reduced, such as by the presence of wood as suggested by the authors elsewhere in the document, then these BSAFs are not only within theoretical possibility, they are the theoretical expectation.

Response

The site-specific laboratory bioaccumulation studies resulted in BSAFs substantially greater than are reported in the literature for field studies. This was summarized in the BERA. In the revised BERA we will expand the basis for comparison among field and laboratory study-derived BSAFs.

Benthic community study

46. The text emphasizes that the benthic community study is the most important line of evidence because it incorporates real world exposures and the actual invertebrate community exposed. While these conceptual arguments are true, the ability of a benthic community study to represent these qualities lies in the ability of the study to detect differences, if they exist. If a study encounters a large degree of variability such that discriminatory power is greatly decreased, then the strength of the benthic community study as a line of evidence is decreased commensurately. It appears that there was tremendous variability encountered (which is not unusual), even in the PAH exposures. Plots included in the Appendix 2 indicate that the range of PAH concentration

measured at SQT 1 overlapped the range of most or all site locations (incidentally, log scales would be a big help in displaying data with large ranges in values). With that kind of heterogeneity within a sampling location, is it really reasonable to expect much discriminatory power in associating exposure with community condition? Further, there was considerable scatter among reference stations, which would further cloud detection of community differences. There is neither discussion nor analysis of the power of the benthic community study. This should be done, and I suspect it will find that the power of this study is low. That is not necessarily the fault of those conducting the survey; it is an all too common problem with benthic community data in general, particularly for sites with heterogeneous substrate and contamination. This needs to be addressed in the discussion of the data and of uncertainties. If the benthic community study has low power, then it is prone to underestimating effects and is in fact a weak line of evidence rather than a strong one.

Response

The reviewer is correct that there is high variability in benthic community data and that is why substantial care was taken to design a study that could distinguish the variability due to the presence of PAHs from that due to other environmental variables.

In the first place a power analysis was conducted on SEH (1998) benthic community data to evaluate the number of samples that would be required to detect differences amongst stations. This is explained in detail in the RI/FS work plan.

A statistical approach involving several conservative hypothesis testing methods including regression, step-wise regression, Analysis of Variance (ANOVA), Analysis of Covariance (ANCOVA) and Multivariate Analysis of Variance (MANOVA), as well as other methods that don't rely on the General Linear Model, including box and whisker plots and cluster analysis, was employed to evaluate these data. This approach is described in detail in the Benthic Community Investigation report (Attachment 3 to Appendix B to the BERA.

As designed and implemented, the benthic community evaluation had sufficient power to differentiate the effects of PAHs from other environmental variables. The results indicated there were no significant effects. See response to Comment #16,

47. On page 6-22 there is discussion of selection as a mitigating circumstance for the benthic community. That is a two-edged sword, and not necessarily a mitigation against risk. Many studies have shown loss of genetic diversity associated with selection for contaminant resistance and, while it is difficult to prove, this observation can easily be extended to argue that selection for contaminant resistance decreases a population=s ability to withstand other stressors. It would be best to just call this issue a draw, inconclusive either way.

Response

The reviewer is apparently suggesting that the gene pool of organisms in the Site area is somehow being diminished by the presence of contaminants. The benthic community assessment provides evidence that this has not yet occurred sufficiently to modify the community structure compared to non-contaminated areas. However, it is a hypothetical possibility that if species populations at the Site were restricted for a long period time to this contaminated area and as a result were reproductively isolated from other populations of the same species there could be genetic drift in the population that might decrease the population's ability to withstand some stressors as this comment suggests.

However, this hypothetical situation is extremely unlikely because:

1) The Site area is extremely small compared to the spatial distribution of the species that inhabit the Site.
2) The reproductive strategies of the species that live there vary widely. There are species that brood their young, species that are broadcast spawners, aquatic insects that reproduce after hatching, etc. Most of these strategies incorporate a mechanism for dispersal of the species, which is probably what has contributed to their evolutionary success in the first place.
3) In addition to having reproductive strategies that facilitate dispersal, physical mechanisms, including long shore currents and storm effects as well as commensal relationships with other species contribute to their dispersal and mixing of the gene pool.

As a result of the above factors, organisms at the Site are not reproductively isolated from nearby populations of the same species and selective pressures for tolerance characteristics that also decrease the genetic "vigor" of the species' metapopulation,(i.e. a set of local populations connected by migrating individuals or in this case dispersed genetic stock) as suggested by USEPA, are likely very low compared to other selective pressures.

Specific Comments

1. **Figure 2-3:** The area of impacted sediments should be presented using total PAHs not just naphthalene.
Response

This was done for convenience because we already had a figure showing distribution of naphthalene and it is approximately the same as for PAHs. We will revise this figure to show the distribution of PAHs.

2. **Tables 3-1 and 3-2 and Tables 5-1, 5-2, 5-3:** When presenting organic carbon normalized data, the organic carbon content should also be provided.
Response

Although we are unaware that this is a convention, and indeed, very few of the scientific journals focusing on environmental chemistry and risk follow this convention, we will be include this information in the revised BERA.

3. **Section 3.6:** Screening of constituents of potential concern (COPCs) based on the 95th percentile upper confidence limit on the mean concentration (95UCL) should not be done. All chemicals whose maximum detected concentration exceeds the screening level should be carried through the BERA risk characterization. The 95UCL should only be used in the risk characterization. In addition, after screening with the maximum concentrations, chemicals that are bioaccumulative should be retained in the BERA (e.g., mercury) even if present below screening levels. Based on the RI/FS Work Plan (Section 4.3.6.2.1.2.1), COPCs from the earlier risk assessments that would be retained as COPCs were benzo(a)pyrene, benzo(a)anthracene, xylenes, ethylbenzene, cyanide, copper, lead, mercury, and zinc. The COPCs evaluated in the BERA should begin with this list and add new COPCs based on the screening.

Response

We don't understand this comment. It says all chemicals whose maximum exceeds the screening level should be carried through the risk characterization and "The 95UCL should only be used in the risk characterization."

The work plan indicated all chemicals would be re-screened and not that previously screened in chemicals would be presumptive COPCs. All in this list except mercury and lead were retained. Although methyl mercury is bioaccumulative it is not a Site COPC and was not quantified here as was in accordance with the approved work plan. Only total mercury was quantified.

Neither lead nor mercury were carried thorough the risk characterization because their 95UCL did not exceed the TEC..

4. **Section 3.8.2:** Provide the total acreage of upland habitat.

Response

As reported in the report, "Characterization Of Wetlands And Terrestrial Habitats" (Appendix E to the BERA, the total upland area of the site is approximately 23 acres. Of this, the two wood/shrub areas comprise approximately 3.8 acres. The open field between the wooded area and Chequamegon Bay is approximately 7.5 acre.

5. **Table 3-9 and Section 3.11.2.2, Page 3-23 and 3-26:** It is not clear why Table 3-9 states that exposure to chemicals by fish via food transfer will not be evaluated quantitatively, but section

3.11.2.2, which discusses measurement endpoints related to Assessment Endpoint #2, includes a comparison of tissue levels of PAHs and estimated VOCs in fish to the No Effects Body Residue (NEBR). Please clarify.

Response

The discussion in Section 3.11.2.2 on using tissue levels in fish as a measurement endpoint was inadvertently left in. This was not used as a measurement endpoint.

6. **Section 3.11.2.3, Page 3-27, second bullet:** In order to reduce the amount of uncertainty in the black duck food chain model, it is recommended that plants are included in the dietary composition of the black duck.

Response

We don't believe this is appropriate because:
1) There are no site-specific data on contaminant levels in plants or appropriate models or uptake factors to estimate bioaccumulation from sediments or surface water to aquatic plants; and
2) There is little, if any, submergent vegetation in the Site area that black ducks could feed upon.

Using a diet of 100% invertebrates is more conservative then using a proportion of their diet and since there was no risk with an assumed diet of 100% invertebrates, apportioning some of a black duck's diet to plants would result in less of a risk

7. **Section 3.11.3:** The receptors of concern (ROCs) for the aquatic habitat include a bat and tree swallow as insectivorous receptors. These species do not ingest sediment. Aquatic-dependent species that ingest sediment while foraging/nesting should also be evaluated as a ROC.
Response
There was no shoreline habitat that would support foraging species. The shoreline is mostly rip rap. We did model sediment uptake with the duck.

8. **Section 5 and Appendix I:** The site-specific BSAFs should be based on the 95UCL concentration not on the geometric mean.

Response

The draft BERA demonstrates that the BSAFs calculated as the geometric mean of the BSAFs determined from the Lumbriculus variegatus bioaccumulation study are an order of magnitude greater than those typically observed in field studies as reported in the literature (See Section 5.2.3 of the draft BERA and Response 40 above). Given the already conservative nature of the BSAFs used in the draft BERA, NSPW does not believe that it is necessary to introduce additional levels of conservatism and uncertainty into the final BERA by calculating tissue concentrations based on

the 95% UCL of BSAFs. However, NSPW will present tissue concentrations based on BSAFs based on the 95% UCLs of BSAFs as an evaluation of uncertainty.

9. **Section 5.1.2, Page 5-8:** This section discusses the possibility that fish might metabolize PAHs to more toxic metabolites. The first sentence in the first full paragraph states that detoxification is the major fate for PAHs should be revised to make it clear that while detoxification MAY occur, it is not the only possible outcome of PAH metabolism in fish.

Response

We will revise this sentence to indicate it is not the only fate. However, the statement is true as originally stated. Detoxification is the ***major*** *fate. Detoxification mechanisms lead to clearance and the efficient clearance of PAHs is the reason that PAH residues are measured in the bile of fishes, not the liver. Only a very small subset of a few very poorly soluble PAHs (i.e., those with a "Bay" region) are even capable of being converted to procarcinogens and subsequently converted to carcinogens, and even with these compounds, the relative yield is very low. See papers from Sikka's laboratory at SUNY- Buffalo (Sikka et al. 1990a,b; Steward et al 1990a, b; Pangrekar et al. 1995, 2003; Willett et al. 2000; Shappell et al. 2003).*

10. **Section 5.1.2.2, Page 5-16:** The statement that claims that numerous studies show that the Critical Body Residue (CBR) provides a better estimate of toxic concentrations than sediment or surface water benchmarks should have references. In addition, this statement may not be correct because fish metabolize PAHs rapidly and thus sediment and surface water concentrations are useful measures of exposure.

Response

Neither bulk sediment nor surface water measurements account for bioavailability, but this essential parameter is inherent in the CBR. McCarty et al. (1985), DiToro and McGrath (2000), and USEPA (2003) have shown that, although the aqueous LC50 concentrations may differ dramatically, the CBR is uniform within a class of organic compounds. The Target Lipid Model, although firmly based in theory is supported by experimental evidence and the CBR developed for PAHs using this model includes the component of metabolism.

The measured body burden is dependent on the net flux of PAHs into the body. This is the balance between uptake and depuration. When exposure conditions are constant, as in sediment pore space, the rate of depuration is largely dependent on the rate of metabolism, which is relatively lower for invertebrates than for fish. Therefore, the CBR is reached more quickly for poorly metabolized chemicals than for PAHs.

11. **Section 5.1.3, Page 5-24, last paragraph:** The statement that resins and asphaltenes are non-toxic is incorrect. The organic compounds that are found in these substances can be released and can be toxic.

Response

Although hypothetically possible it is highly unlikely that resins and asphaltenes would release lower molecular weight PAHs or alkanes under the conditions experienced in the sediment where minimal weathering has occurred. We could find no evidence of this phenomenon in the peer-reviewed literature. Please provide references so that we can include this caveat.

12. **Section 5.1.3.1, Page 5-25, first paragraph:** The document states that worms will be used as surrogates for invertivorous (invertebrate eating) wildlife. This does not make sense from an ecological or modeling standpoint as worms are in different trophic levels from birds or mammals that consume invertebrates. The statement should be corrected.

Response

This paragraph should have indicated that worm body residues would be used as surrogates for the invertebrate organisms fed upon by invertivorous birds. This will be revised.

13. **Section 5.1.3.2, Page 5-26, Mammals:** The use of naphthalene to represent all PAHs is not wholly acceptable. This compound may not be the most toxic; therefore an analysis using naphthalene to represent all PAHs may not be the most conservative technique for analyzing effects of PAHs on mammals. Secondly, there are additive effects of PAHs in mixtures which may not be reflected in this technique.

Response

We used naphthalene to conservatively represent the toxicity of all PAHs because 1) low molecular weight PAHs predominated the fish and worm tissue data and food consumption is the major exposure pathway to vertebrates, and 2) the only relevant toxicity data for high molecular weight PAHs come from studies with weathered crude oil or coal tars. The TRV for naphthalene was lower by a factor of 3.6-times, hence more conservative.

ATSDR (2005) recently reported the results of an exhaustive toxicity review for naphthalene and methylnaphthalene. For chronic exposures (13 weeks), they reported NOAELs of 400 and 200 mg/kg BW/day for rat and mouse reproduction, respectively). Therefore, our proposed TRV of 129 mg/kg BW/day for mammals for naphthalene is conservative.

In response to the second comment, additivity is accounted for in the dose, not in the TRV. In the BERA the sum of all Class I narcotic chemicals including all PAHs and VOCs were used as the dose. This is consistent with both USEPA guidance (USEPA 2003) and the Target Lipid Model (DiToro, et al. 2005) as explained in the BERA.

14. **Section 5.1.3 and 5.1.4:** The proposed TRVs need to be re-evaluated. Use USEPA Region 9 BTAG low and high TRVs as the primary source of TRVs. Secondary sources can be consulted if the COPC is not listed by Region 9 BTAG. When using the Ecological Soil Screening

Level toxicity data, use both the NOAEL and LOAEL; do no extrapolate a LOAEL using a conversion factor of 5. This also applies to the other studies; a LOAEL from the literature should be selected and used, preferably from the same study if available, rather than a conversion factor of 5. A conversion factor is appropriate when a LOAEL is not available in the literature.

Response

The Region IX BTAG TRVs are really secondary sources. Region IX EPA has cautiously avoided association with the so-called Region IX BTAG values and has continuously referred to them as the EFA West (Engineering Field Activity West) TRVs because they were funded by the Navy. Therefore, the Region IX BTAG low and high TRVs referred to above are only available from the California Department of Toxic Substances (DTSC) website. Furthermore, these values:

1) Exist in at least two forms, one published on the DTSC web-site, and one agreed upon in the Vandenberg Air Force Base dispute resolution, wherein an expert panel from the University of California at Davis rejected a number of the values still provided on the DTSC website;

2) were based on only a cursory review of the available literature, and where the ECO-SSL group has conducted a review, all of the Region IX TRVs have been overturned and increased;

3) use toxicity endpoints such as hair loss and carcinogenicity rather than reproduction or population stability as is recommended in USEPA (1997) risk assessment guidance as well as the ECO-SSLs; and

4) are screening values that by California Guidance are replaced by values from the primary literature when conducting a Predictive Ecological Risk Assessment, the California equivalent to USEPA's Baseline Ecological Risk Assessment.

As explained in the risk assessment, there is a scientifically valid, although highly conservative, reason to extrapolate to a LOAEL using a factor of 5. We only extrapolated values when alternative, high quality, data were unavailable. Simply picking the next higher value tested above the NOAEL is not an appropriate method of selecting the LOAEL, because there may be a large data gap in which effects could occur. In practice, since no risk was found to birds or mammals at the NOAEL, use of the LOAEL was not necessary.

15. **Section 5.1.4.2, pg. 5-28:** The document states that the use of BSAF (biota-to-sediment-accumulation-factors) is "an unreliable way to evaluate the potential for adverse effects." While there is a degree of uncertainty inherent in modeling exposure, this does not necessarily make the technique unreliable. Second, the statement that "Most studies have shown that the major exposure pathway for fish to metals..." does not have references, and is therefore unsubstantiated. References should be included. Third, the last sentence in the section states that "Since there were no exceedances of screening benchmarks for metals in surface water, there is little reason to believe that metals would be elevated significantly above normal levels in Site fish." This

statement is unsubstantiated (at least at this point in the risk assessment). Some metals, particularly mercury, bioaccumulate, which is why food web exposure models are done. Bioaccumulation potential needs to be discussed here.

Response

Perhaps "highly uncertain" rather than "unreliable" should have been used to describe the use of BSAFs to evaluate the potential for adverse effects.

There are only two exposure pathways for metal uptake, and both are discussed. It is recognized in the USEPA National Ambient water Quality Criteria that dissolved metals are the dominant bioavailable form of metals. The statement regarding this major uptake pathway is supported by Clearwater et al. (2002), among others, who reported that while the gill has no ability to regulate the uptake of metals, the liver regulates uptake from the diet. This is reflected in biominimization of metals at progressively higher trophic levels although exposed to the same aqueous environment. Hepatic regulation is the basis for homeostasis of essential micronutrients such as copper and zinc, but also for the detoxification and depuration of metals such as cadmium and mercury. Methylmercury is the only metal that biomagnifies in a toxic form through the food chain. In accordance with the approved RI/FS Work Plan, metals were not measured in surface water, however, there is no evidence to suggest that mercury should be elevated in Site surface waters
See also: Clearwater et al. (2000) [copper], Spry et al. 1988) [Zinc] , Bury et al.(2003) [essential micronutrients].

16. **Section 5.2.2:** Surface water intake should be quantified for all higher level receptors evaluated through food chain modeling.
Response

There were only occasional low level detections of benzene, ethylbenzene, toluene and naphthalene in the filtered fraction of Site surface water and none of these detections exceeded screening criteria. No other VOCs or PAHs were detected. For this reason it was decided not to model the surface water exposure pathway.

17. **Section 5.6 and Appendix I:** What data set was used to develop soil and sediment EPCs? Sediments far from shore should not be included in EPCs used for quantifying sediment ingestion by higher level ecological receptors, as these organisms are not expected to be exposed to sediments far from shore. Was the SEH data included in the sediment EPC dataset? Data from what soil and sediment depths was used? Why do the BERA and HHRA soil data sets for Kreher Park differ?

Response

All data for sediment, including historical SEH data, and soil that met the criteria for an exposure pathway was used in the BERA. As explained and illustrated in Appendix A to the BERA, sediment

samples were primarily those in the 0-6" interval although samples from the 0-24"interval were used in areas where there was not adequate coverage of 0-6" samples. Soil samples from the 0-12" interval were used as the basis for the soil exposure point concentration. This was not explicitly stated in Appendix A but will be revised in the final BERA.

The only wildlife ROC where incidental ingestion of sediment was considered an exposure pathway was the black duck. Black duck are known to dive 2 to 3.8 meters (6 to 12 feet) for food (Brodsky and Weatherhead 1985), so it was assumed that a black duck could feed throughout the Site.

The BERA excluded two soil samples which were used in the HHRA. In one instance it was because the soil interval was uncertain, in the other because it indicated it was a "seep boring". We will reconcile soil sample selection criteria with the HHRA so they are consistent.

The BERA also defaulted to 95%UCL based upon a bootstrap or a jackknife if the EPC data were not normally distributed, while the HHRA used the 95% Chebyshev estimation that ProUCL recommended. This will be re-evaluated in the revised BERA.

18. **Table 5-12:** The estimated tissue concentrations are presented as wet weight in this table, and dry weight in the Appendix I table. Ensure that dry weight tissue concentrations are used in the intake calculations as the ingestion rates used are on a dry weight basis.

Response

Table 5-12 presents estimated concentrations of PAHs and VOCs in benthic invertebrate tissue on a µmol/g lipid basis for comparisons with no effect body residue (NEBR) values. In the calculations, PAH concentrations were first estimated as wet weight tissue concentrations (mg PAH/kg tissue, wet weight) based on the wet weight fraction of lipids. The wet weight tissue concentrations (mg PAH/kg tissue, wet weight) were then converted to µmol/g lipid. Because the final concentrations were ultimately expressed on a mass lipid basis (µmol/g lipid), the same concentration would have been achieved if dry weight tissue concentrations (mg PAH/kg tissue, dry weight) were initially calculated using the dry weight fraction of lipids.

URS recognizes that Table 5-12 may be simplified by calculating benthic invertebrate tissue concentrations on a µmol/g lipid basis directly from the BSAFs (kg organic carbon/kg lipid) obtained from the bioaccumulation study. This modification will not change the tissue concentrations (µmol/g lipid) that were compared to the NEBR in the BERA, but it will eliminate the confusion associated with wet weight versus dry weight tissue concentrations (mg PAH/kg tissue). Table 5-12 will be modified in the final BERA to reflect this change.

19. **Table 6-1:** Individual sample locations (not an average concentration) should be compared to TECs/ PECs.

Response

This is more appropriately done in the nature and extent part of the RI not in the BERA.

20. **Section 6.2.1, Page 6-2 and Appendix B:** PEC-quotients should be developed for each individual sediment sample to evaluate cumulative impacts from metals and PAHs.

Response

We disagree because:

1) There is no ecological exposure to only one sample;
2) Doing it would not provide any more insight into so-called cumulative effects than what was done in the BERA.

PAHs and VOCs found in Site sediments all have the same mechanism of action, i.e. narcosis, but this mechanism of action is different than that of the few other Site COPCs. Therefore, cumulative effects, if any, are unknown and cannot be estimated using the current approaches.

21. **Section 6.2.1.6, Page 6-5, second full paragraph:** Is there evidence to substantiate the claim that the levels of site chemicals are higher in the top two or three inches of sediments than in the top six inches? If there are data supporting this claim, they should be referenced and discussed. Otherwise this statement is conjecture. This comment also applies to Appendix B, Section 4, Page 4-3, first paragraph.

Response

The reviewer misinterpreted what the BERA stated. It was suggested that levels in the top two inches are lower than deeper in the 0-6" interval and that would explain absence of effects to benthos.

22. **Section 6.2.13, Page 6-13, last paragraph:** This paragraph discusses evolutionary adaptations and relative susceptibilities to chemical exposure by organisms at different trophic levels. It is not necessarily true that "lower" aquatic organisms would not be adversely affected by chemical exposure if it were shown that "higher" organisms were not significantly affected by the same levels of chemical contaminants. In some cases, some "lower" groups are more susceptible than "higher" groups. This paragraph should reflect this possibility.

Response

The BERA will be revised to reflect this possibility, however, the point that was being made was related to the functionality of "lower" level communities in the ecosystem. While conditions may

affect "groups" or populations of "lower" level organisms, the absence of effects in "higher" level organisms should be evidence that the functionality of the lower level organisms has not been materially impaired.

23. **Table 6-3, Page 6-18:** The statement that LOAELs are more reliable than NOAELs is not necessarily correct. The use of LOAELs versus NOAELs depends on the situation. The reliability of NOAELs is not in question. The statement should be reworded.
Response

We disagree with the reviewer that the reliability of NOAELs is not in question. NOAELs are merely doses that caused no effects; they cannot indicate risk because they are not necessarily thresholds above which effects are found. For instance, a dose of zero is a NOAEL. LOAELs, on the other hand, are doses that have been demonstrated to result in an effect. Van der Hoven (1997) summarized some of the problems associated with NOAELs:

1) they are based on wrong usage of hypothesis testing- the acceptance of the null hypothesis;
2) they depend on the accuracy of the test- when sample error is small, the probability of finding a lower NOAEL increases;
3)they depend on sample size- the larger the sample, the higher the probability of finding a lower NOAEL; and
4) they depend on the chosen significance level.

Many authors recommend low effects concentrations as alternatives to NOAELs and LOAELs. NOAELs are also commonly used only to evaluate risks to listed species, while LOAELs are used to evaluate risks to non-listed species. In aquatic toxicity testing, the geometric mean of the LOEC and NOEC is considered the concentration in which population stability is not affected. See Chapman et al (1996), Kooijman (1996), Moore and Caux (1997), Crane and Newman (2000), and USEPA Comment10, above.

24. **Section 6.3.3.3, Page 6-25:** The three statements are misleading. Bioconcentration factors and biota sediment accumulation factors, while limited, are available in the literature for inorganics, and literature-based BSAFs were used in the document (Appendix F). Copper and selenium are known to occur at MGP sites, and both are known to be toxic to mammals and birds at concentrations slightly above nutritional requirements. While uptake into fish and invertebrate tissue from sediments or surface water may not be able to be quantified for COPCs (an uncertainty that potentially will underestimate risk), risks to aquatic receptors who forage along the shore and incidentally consume sediments contaminated with metals needs to be evaluated in this BERA.

Response

The fact that BSAFs are available for metals does not mean they are accurate or can be applied with any degree of confidence. For example, copper is an essential micronutrient that is efficiently regulated by birds and mammals. Therefore, copper is known to be non-toxic at concentrations

"slightly above nutritional requirements". In some instances selenium is toxic at concentrations slightly above nutritional requirements, but this depends largely upon site-specific factors. Selenium toxicity to birds or mammals has only been shown at elevated concentrations associated with agricultural runoff and high evaporation rates in certain areas of the Great Plains (USGS 1999) and the far West. In the Great Plains area, selenosis or alkali disease is only found in domestic ruminants (Moxon 1937), not wild ruminants such as antelope (Raisbeck et al. 1996). In the far West, selenium may be toxic to birds but again, only at extraordinarily high exposure rates associated with agricultural drainage and high evaporation. These conditions do not exist at the site.

Finally, there is no shoreline for wildlife to forage on at the site. The shoreline is covered with rip-rap.

25. **Appendix B:** Attachment 2. For the bioassay results, also present the following information in the summary tables:

Sediment Quality Metrics
Chemical/Physical
TOC (%)
AVS (umol/g)
% Fines (<63 um)
Mercury (ppm)
Empirical
PEC-Q metals
PEC-Q PAH
Mean PEC-Q
LRM Pmax
Mechanistic
ESBTU
(SEM-AVS)/foc

Toxicity Metrics
Hyalella azteca
Survival
Growth
Temperature □C (Day 0)
pH (Day 0)
Ammonia (total mg/l NH_3 - Day 0)
Ammonia (unionized mg/l NH_3)

Chironomus dilutus
Survival
Growth
Temperature □C (Day 0)
pH (Day 0)
Ammonia (total mg/l NH_3 - Day 0)
Ammonia (unionized mg/l NH_3)
Pimephales promelas

Survival
Growth
Temperature □C (Day 0)
pH (Day 0)
Ammonia (total mg/l NH_3 - Day 0)
Ammonia (unionized mg/l NH_3)

Response

All of this information, other than the various empirical and mechanistic quotients can be found in Appendices 2 to 20 of Attachment 2 to Appendix B. Since we conducted a sediment quality Triad and have site-specific data, we don't believe it is necessary to go any further than comparing our bulk sediment chemistry results to the Wisconsin Sediment Quality Benchmarks. However, the data are there for the reviewer to calculate these quotients.

26. **Appendix B, Section 4, Page 4-2, 1st paragraph, last sentence:** This statement is misleading and should be rewritten. The LSRI report noted that the survival in the performance control samples for *C. dilutus* and the minnow met acceptability criterion. The minnow showed significant effects on growth at SQT1 and under UV light there was significant reduction in survival at SQT1 and SQT7. Overall, it appears that the reference stations have been impacted and comparisons to performance control samples should be given more weight than comparison to the reference samples.

Response

We agree with the reviewer that the reference stations appear to have been toxic. However, we believe that this means that the Site stations were most likely affected by non-Site-related conditions not that the Site data should be compared against performance controls. This is the reason reference stations are evaluated in bioassay programs.

27. **Appendix B, Table 3-1:** Why was the organic carbon content of SQT3 (sand) so high (40%)? This percentage is similar to the highest concentrations found in the wood stations (42%). Should this location be excluded from the sand samples and included as a wood station?

Response
Station SQT3 was designated to be a Sand Station and after collecting some trial samples during the reconnaissance it was sampled as a sand sample. After the lab results were received, it was clear that actual samples had higher levels of organic carbon than was planned. The five bulk sediment replicates ranged from 17 to over 40% TOC.

The fact that this station is called a Sand Station but really has high organic carbon is factored into the interpretation of the bioassay.

28. **Appendix B, Attachment 2:** NOEC and LOEC values presented in the tables and text do not match. Values in Tables 1-6, 3-1, and 5-2 need to be verified with the text. For instance, text on page 3-1 does not match numbers in Table 3.1. AAverage NOECs ranged from 735 to 8031. Should 8031 be 7257 µg PAH/gOC as presented in Table 3.1? Also change 139.5 µg/g to 135.1. In last sentence, should 3,396 be changed to 3,996.

Response

The reviewer is correct that the value of 8,031 ug PAH/gOC is incorrect. This will be corrected in the revised document. The values in Tables 1.6 and 3.1 don't agree because Table 1.6 refers to URS tests conducted in 2005, while Table 3.1 refers to the average of URS and SEH results for tests conducted in 2001 and 2005. The text on page 3-1 explains these differences.

29. **Appendix B, Attachment 2, Section 1.4, Page 1-6:** The statement that no NOECs or LOECs are proposed for *H. azteca* contradicts the tables and discussion in the main text (Section 5.1.2.2) where NOECs and LOECs are presented, discussed, and (presumably) incorporated into the average and range of NOECs and LOECs. In addition, the statement on 3-1 (Appendix B, Attachment 2, Section 3) that describes the range of average NOECs (735 to 8031 ug PAH/gOC) also is in disagreement with the statement in section 5.1.2.2, pg 5-18 that describes the range of average NOECs (735 to 7257 ug PAH/gOC). The values should be corrected to be consistent.

Response

This section described tests conducted by URS in 2005 and 2006. On page 1-6 we proposed a NOEC of 4536 µg PAH/gOC for H. azteca. The statement on page 1-6 referred to above described the URS Chironomus bioassays, not H. azteca bioassays. Or statement was "Due to the high reference site mortalities for the midge, no NOEC or LOEC values are proposed for this species." We ultimately proposed NOECs and LOECs for Chironomus based solely on the SEH 2001 bioassays described on pages 2.2 and 5.2. The other values cited above will be corrected in the revised BERA. See Response 28 with regard to the value of 8,031 µg PAH/gOC

30. **Appendix C and Appendix I:** The fish tissue concentration in Appendix C is on a wet weight basis; the 95UCL tissue concentration presented in Table I-3 is on a dry weight basis. Please clarify which is correct. A summary table of fish tissue data should be provided in the text similar to the summary tables for sediment and soil presented in Appendix A. How was the 95% UCL concentration of total PAHs in fish calculated?

Response

Fish tissue analyzed for the BERA and HHRA investigations were reported by the laboratory on a wet weight basis, and therefore were presented on a wet weight basis in Appendix C, the Fish Tissue Investigation report. Because all parameters in the wildlife dose rate models in Appendix I were expressed on a dry weight basis, the measured wet weight fish tissue concentrations presented in Appendix C were converted to dry weight assuming an average moisture content of

80% in fish. A footnote will be added to Table I-3, to clarify the origin of the wet weight fish data and its conversion to a dry weight exposure point concentration for its application in dose rate modeling. A table will also be provided to Appendix C of the final BERA that will summarize the results of the total PAH analyses in fish tissue.

*The 95% UCL concentration of total PAHs in fish should have been calculated as the 95% UCL of total PAH concentrations measured in whole body smelt, smallmouth bass, and brown bullhead collected from the Site (expressed on a dry weight basis as described above). However, the 95% UCL concentration of total PAHs presented in the draft BERA was inadvertently based on the 95% UCL of total PAH concentrations measured in whole body samples of these species collected from the Site **and** reference areas. In the final BERA, the 95% UCL concentration of total PAHs will be based only on whole body measurements of smelt, bass, and bullhead collected from the Site area. The change in exposure point concentrations for fish will be reflected in the dose rate models for piscivores in the final BERA.*

31. **Table F-1:** The 90th percentile percent soil in the diet of a dove of 13.9% (USEPA, 2005) should be applied for the blackbird.

Response

We believe the proposed 90th percentile soil ingestion rate is an overly conservative parameter for the red-winged blackbird in a baseline assessment. The 90th percentile soil ingestion rate is utilized in the development of Eco-SSLs, which are intended as a screening-level tool to 'identify concentrations of contaminants in soil that may present a risk from those concentrations that clearly do not' (USEPA 2005). In the Eco-SSL guidance, USEPA considers the use of the 90th percentile soil ingestion rate to be a highly conservative parameter in its models.

The soil ingestion rate of 8.8% used in the draft BERA is consistent with and, in most cases, more conservative than soil ingestion rates for wildlife with similar dietary habits. Beyer et al. (1994) provides soil ingestion estimates of 10.4 and 9.3% for American woodcock and wild turkey, respectively. Woodcock probe into the soil and forage almost exclusively on earthworms and other ground-dwelling invertebrates, resulting in greater incidental ingestion of soil relative to wildlife whose diet consists of at least a portion of plant material. Assuming a dietary composition of 20% invertebrates and 80% plant material, Sample and Suter (1994) estimated a soil ingestion rate of 2.08% for the American robin based on American woodcock data. Applying the same approach as Sample and Suter (1994) and assuming the draft BERA estimates the dietary composition of red-winged blackbird to be 50% vegetation and 50% plant material, the soil ingestion of red-winged blackbird is estimated to be 5.3%. Based on these comparisons, NSPW concludes that the soil ingestion rate of 8.8% for the red-winged blackbird is sufficiently conservative for the BERA.

32. **Table F-2:** The regression model in USEPA (2005) should be used to estimate soil invertebrate concentrations for cadmium. Provide soil BAFs for dibenzo(a,h)anthracene. EPA 2005 provides a regression model for total PAH uptake into foliage. The BSAFs for sediment invertebrates and fish should be based on 95UCL tissue concentrations not the geometric mean.

Response

We believe that the regression model used in the draft BERA to estimate soil invertebrate uptake of cadmium provided by Neuhauser et al. (1995) is more appropriate than the model presented in USEPA (2005) derived from Sample et al. (1999). In developing of bioaccumulation models, Sample et al. (1999) compare their models with models previously published in the literature. In the case of cadmium uptake by soil invertebrates, Sample et al. (1999) indicate that 'the model that best estimated Cd concentrations in earthworms was from Neuhauser et al. (1995) and not the current study'. Based on this appraisal, the cadmium uptake model presented by Neuhauser et al. (1995) will remain in the final BERA.

Dibenzo(a,h)anthracene was not detected in soil samples. Therefore, terrestrial receptor exposure to dibenzo(a,h)anthracene was not included in the dose calculated for total PAHs.

The concentration of PAHs in plants estimated by the USEPA (2005) regression model was greater than the total PAH tissue concentration estimated in the draft BERA. Based on the 95% UCL concentration for total PAHs in soil samples from the site (49.34 mg/kg) and the regression model for rinsed plant foliage (the model for unrinsed foliage is not significant), the estimated concentration of PAHs in plant tissue is 6.96 mg/kg (dry weight). The draft BERA calculated a total PAH concentration in plants of 23.0 mg/kg (dry weight) based on the summed concentrations of individual PAHs estimated using regression models for individual PAHs (USEPA 2005). Given that HQs for herbivores exposed to total PAHs were substantially less than 1.0 based on a more conservative estimate of total PAH, NSPW proposes that the estimation of total PAH concentrations used in the draft BERA remain in the final BERA.

As described in the response to Comment #8, NSPW believes that the BSAFs calculated as the geometric mean of the BSAFs determined from the Lumbriculus variegatus bioaccumulation study are adequately conservative for estimating concentrations of PAHs in benthic invertebrate tissues. NSPW does not believe that it is necessary to introduce additional levels of conservatism and uncertainty into the final BERA by calculating tissue concentrations based on the 95% UCL of BSAFs. However, NSPW will present tissue concentrations based on BSAFs based on the 95% UCLs of BSAFs as an evaluation of uncertainty.

33. **Appendix I:** Dose tables should provide all intake factors used for the receptor including the BSAFs, prey intake rates, soil intake rates, and body weight. For the osprey, cormorant, and mink present the measured fish tissue concentrations.

Response

The information requested in Comment #33 is provided in Tables I-1 through I-3 of Appendix I, which precede the dose Tables I-5 through I-12. This information is included in dose tables using a "look up" link in Excel ™. The level of detail presented in Tables I-1 through I-3 is intended to provide the utmost transparency in the calculation of wildlife doses. NSPW does not believe it is practical to include this level of detail in the dose tables.

Typos and Editorial Errors

34. **Section 2.3, Page 2-2, first sentence:** Sentence should read: "The Site is located at the top of a ravine on the shore of Chequamegon Bay."

Response

This change was made.

35. **Section 3.2.2, Page 3-2, second paragraph:** The statement: "SEH (2002) reported that the levels..." is not relevant to an Ecological Risk Assessment. Delete this statement.

Response

This sentence will be deleted.

36. **Section 3.8.1.2, Page 3-13, last paragraph:** Sentence should read: "...the list of those species frequenting the Site waters..."

Response

This change was made.

37. **Section 5.1.2.2, Page 5-17, second full paragraph:**
a) Sentence should read: "The results..."
b) Section on Hyallela; Sentence should read: "Furthermore, the mortality was consistent..." pg. 5-18, Section on Chironomus, last paragraph; and Tables 5-1 to 5-3
c) This paragraph needs clarification and rewriting. The statement beginning "When the NOECs and LOECs..." needs proper punctuation and clarification. Are the NOECs/LOECs for each species averaged separately or lumped together? The statement should read that the NOECs for each species are averaged between the 2001 and 2005/6 results, correct?
d) The average LOEC for H. azteca is 422.8 ug/g, not 453.3 ug/g. Therefore, the range of average LOECs is 79.9 ug/g to 422.8 ug/g, not 208.3 ug/g. The average LOEC for total PAHs per gOC for H. azteca is 5463 ug/gOC, not 11494 ug/gOC.
e) pg. 5-20, first full paragraph: Sentence should read: "The unionid snail, also an epibenthic

species, was absent from all of the Site..."

Response

These changes and clarifications were made

38. **Section 5.1.4.3, Page 5-30, Selenium:** The sentence: "However, both laboratory and field studies..." should read "accumulate egg selenium concentrations much greater than 3 mg Se/kg dry weight without adverse effects on reproduction." (delete "are").

Response

These changes were made.

39. **Section 6.2.1.2, Page 6-3:** The sentence "At the other stations..." is not clear. Please rewrite for clarity. Also, use of contractions is not appropriate. (See also Section 6.2.1.5, pg. 6-4).

Response

These changes were made.

40. **Section 6.2.1.4, Page 6-3:** The first sentence should read "...indicated that there were significant effects..."

Response

This change was made.

41. **Section 6.2.2, Page 6-6, first numbered bullet:** Sentence should read: "Compare ~~of~~ concentrations of Site-related..." (delete "of").

Response

This change was made.

42. **Section 6.2.14, Page 6-14, second paragraph:** Second to last sentence should read "thick, dull iridescence with brown streaks about..."

Response
This change was made.

43. **Section 6.3.2, Page 6-18, last paragraph:** Sentence should read: "The information presented, while as complete and accurate as possible, may have missed...". It is extremely

unlikely that in any scientific study, information will be complete, especially in a Superfund site.

Response

This change will be made.

References

ATSDR. 2005. Toxicological profile for naphthalene, 1-methylnaphthanene, and 2-methylnaphthalene. Agency for Toxic substances and Disease Registry, U.S. Department of Health and Human Services, Washington, D.C.

Bell, H., K. Liber, D. Call, and G. Ankley. 2004. Evaluation of bioaccumulation and photo-induced toxicity of fluoranthene in larval and adult-stages of Chironomus tentans. Arch. Env. Contam. Toxicol. 47: 297-303.

Beyer, W.N., Connor, E.E., and S. Gerould. 1994. Estimates of soil ingestion by wildlife. Journal of Wildlife Management 58(2):375-382.

Brodsky, L.M., and P.J. Weatherhead. 1985. Diving by wintering Black Ducks: an assessment of atypical foraging. Wildfowl 36:72-76.

Bury, N., P. Walker, and C. Glover. 2003. Nutritive metal uptake in teleost fish. J. Exp. Biol. 206: 11-23.

Chapman, P.M., R. Caldwell, and P.F. Chapman. 1996. A warning: NOECs are inappropriate for regulatory use. Env. Toxicol. Chem. 15: 77-79.

Clearwater, S., S. Baskin, C.M. Wood, and D. McDonald. 2000. Gastrointestinal uptake and distribution of copper in rainbow trout. J. Exp. Biol. 203: 2433-2466. 2000.

Clearwater, S., A. Farag, and J. Meyer. 2002. Bioavailability and toxicity of dietborne copper and zinc in fish. Comp. Biochem. Physiol. 312C: 269-313.

Cohen Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Hillsdale, NJ, Lawrence Erlbaum Associates, Publishers.

Crane., M and M. Newman. 2000. What level of effect is a no observed effect?. Env. Toxicol. Chem. 19: 516-519.

De Haas, E, C. Wagner, A. Koelmans, M. Kraak, and W. Admiral. 2006. Habitat selection by chironomid larvae: fast growth requires fast food. J. Anim. Ecol. 148- 155.

Diamond, S., N. Milroy, V. Mattson, L. Heinis, and D. Mount. Draft. Photo-activated toxicity in benthic organisms collected from PAH contaminated sites. Cited by Chapman et al. 2002.

DiToro, D. and J. McGrath. 2000. Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. II. Mixtures and sediments. Env. Toxicol. Chem. 19: 1971-1982.

DiToro, D., J. McGrath, and D. Hansen. 2000. Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. I. Water and tissue. Env. Toxicol. Chem. 19: 1951-1970.

Driscoll, S., and P. Landrum. 1997. A comparison of equilibrium partitioning and critical body residue approaches for predicting toxicity of sediment-associated fluoranthene to freshwater amphipods. Env. Toxicol. Chem. 16: 2179-2186.

Driscoll, S., G. Harley, and P. Landrum. 1997. Accumulation and toxicokinetics of fluoranthene in sediment bioassays with freshwater amphipods. Env. Toxicol. Chem. 16: 742-753.

Edgar, W. and P. Meadows. 1969. Case construction, movement, spatial distribution and substrate selection in the larva of *Chironomus riparius* Meigen. J. Exp. Biol. 50: 247-253.

Erickson, R. 2006. Letter from Russ Erickson (EPA) to Dave Mount (EPA), dated 25 August 2006.

Harkey, G., S. Driscoll, and P. Lamdrum. 1995. Bioavailability of polycyclic aromatic hydrocarbons from a historically contaminated sediment core. Env. Toxicol. Chem. 14: 1551-1560.

Hatch, A. and G. Burton. 1999. Photo-induced toxicity of PAHs to Hyalella azteca and Chironomus tentans: effects of mixtures and behavior. Env. Pollut. 106: 157-167.

Hawthorne, S. and J. Kreitinger. 2003. Predicting the bioavailability of PAHs from MGP soils and sediments using mild supercritical CO_2 extraction. Abstract. www-esd.lbl.gov/CEB/doe_perf_conf/pdf/Collection-final.PDF

Hawthorne, S., C. Grabanski, D. Miller, and J. Kreitinger. 2005. Solid-phase microextraction measurement of parent and alkyl polycyclic aromatic hydrocarbons in milliliter pore water samples and determination of K_{DOC} values. Env. Sci. Technol. 39: 2795-2803.

Ingersoll, C. 1998. Methods for assessing bioaccumulation of sediment-associated contaminants with freshwater invertebrates. Proceedings, National Sediment Bioaccumulation Conference, EPA 823-R-98-002.

Ingersoll, C., P. Haverland, E. Brunson, T. Canfield, F. Dwyer, C. Henke, N. Kemble, D. Mount, and R. Fox. 1996. Calculation and evaluation of sediment effect concentrations for the amphipod *Hyalella azteca* and the midge *Chironomus riparius*. J. Great Lakes Res. 22: 602-623.

Kooijman, S. 1996. An alternative for NOEC exists, but the standard model has to be abandoned first. Oikos 75: 310-316.

Lee, J-H, P. Landrum, J. Field, and C-H Koh. 2001. Application of a Σpolycyclic aromatic hydrocarbon model and a logistic regression model to sediment toxicity data based on a species-specific, water-only LC50 toxic unit for *Hyalella azteca*. Env. Toxicol. Chem. 20: 2102-2113.

Lee, J-H., P. Landrum, and C-H Koh. 2002. Toxicokinetics and time-dependent PAH toxicity in the amphipod *Hyalella azteca*. Env. Sci. technol. 36: 3124-3130.

McCarty, L., P. Hodson, G. Craig, and K. Kaiser. 1985. The use of structure-activity relationships to predict the acute and chronic toxicity of organic chemicals to fish. Env. Toxicol. Chem. 4: 595-606.

McDonald, B. and P. Chapman. 2002. PAH phototoxicity- an ecologically irrelevant phenomenon? Mar. Pollut. Bull. 44: 1321-1326.

Moore, D. and P-Y Caux. 1997. Estimating low toxic effects. Env. Toxicol. Chem. 16: 794-801.

Moxon, A. 1937. Alkali disease or selenium poisoning. S. Dak. Agric. Exp. Sta. Tech. Bull. No. 311, 91 p.

Newfields. 2006. Environmental forensic investigation of tar residues at the Ashland Lakefront site in Ashland, WI.. January 2006.

Norberg-King, Teresa J., Sibley, Paul K., Burton, G. Allen, Ingersoll, Christopher G., Kemble, Nile E., Ireland, Scott, Mount, David R., Rowland, Carolyn D. Interlaboratory Evaluation of *Hyalella azteca* and *Chironomus tentans* Short-Term And Long-Term Sediment Toxicity Tests. Environmental Toxicology and Chemistry 2006 25: 2662-2674

Neuhauser, E., Z. Cukic, M. Malecki, R. Loehr, and P. Purkin. 1995. Bioconcentration and biokinetics of heavy metals in the earthworm. Env. Pollut. 89: 293- 301.

Ozretich, R., S. Ferraro, J. Lambertson, and F. Cole. 2000. Test of Σpolycyclic aromatic hydrocarbon model at a creosote-contaminated site, Elliott Bay, Washington, USA. Env. Toxicol. Chem. 19: 2378-2389.

Pangrekar, J., C. Kandaswami, S. Kumar, and H. Sikka. 1995. Comparative metabolism of benzo(a)pyrene, chrysene and phenanthrene by brown bullhead liver microsomes. Mar. Env. Res. 39: 51-55.

Pangrekar, J., P. Kole, S. Honey, S. Kumar, and H. Sikka. 2003. Metabolism of chrysene by brown bullhead liver microsomes. Toxicol. Sci. 71: 67-73.

Raisbeck, M., D. O'Toole, R. Schamber, E. Belden, and L. Robinson. 1996. Toxicologic evaluation of a high-selenium hay diet in captive pronghorn antelope (Antilocrapa amerinana). J. Wldlf. Dis. 32: 9-16.

Rothstein, Hannah., Jacob Cohen, David Schoenfeld and Jesse Berlin. 2000. SamplePower release 2.0. Chicago, SPSS Inc.

Sample, B. and G. Suter. 1994. Estimating exposure of terrestrial wildlife to contaminants. ES/ER/TM-125. Oak Ridge national Laboratory, Oak Ridge TN.

Sample, B., G. Suter, J. Baeuchamp, and R. Efroymson. 1999. Literature-derived bioaccumulation models for earthworms: development and validation. Env. Toxicol. Chem. 18: 2110-2120.

Shappell, N., U. Carlino-MacDonald, S. Amin, S. Kumar, and H. Sikka. 2003. Comparative metabolism of chrysene and 5-methylchrysene by rat and rainbow trout liver microsomes. Toxicol Let. 72: 260-266.

Sikka, H., J. Rutkowski, and C. Kandaswami. 1990a. Comparative metabolism of benzo(a)pyrene by liver microsomes from brown bullhead and carp. Aquat. Toxicol. 16: 101-111.

Sikka, H., J. Rutkowski, C. Kandaswami, S. Kumar, K. earley, and R. Gupta. 1990b. Formation and persistence of DNA adducts in the liver of brown bullheads exposed to benzo(a)pyrene. Cancer Lett. 49: 81-87.

Spry, D., P. Hodson, and C.M. Wood. 1987. Relative contributions of dietary and waterborne zinc in rainbow trout, Salmo gairdneri. Ca. J. Fish. Aquat. Sci. 45: 32-41.

Steward, A. C. Kandaswami, S. Chidambaram, C. Ziper, J. Rutkowski, and H. Sikka. 1990a. Disposition and metabolic fate of benzo(a)pyrene in the brown bullhead. Env. Toxicol. Chem. 9: 1503-1512.

Steward, A., J. Zaleski, and H. Sikka. 1990b. Metabolsim of benzo(a)pyrene and (-)-trans-benzo(a)pyrene-7,8-dihydrodiol by freshly isolated hepatocytes of brown bullheads. Chem. Biol. Inetr. 7: 119-138.

Swartz, R., S. Ferraro, J. Lambertson, R. Ozretich, B. Boese, D. Schults, M. Behrenfeld, and G. Ankley. 1997. Photoactivation abd toxicity of mixtures of polycyclic aromatic hydrocarbons in marine sediment. Env. Toxicol. Chem. 16: 2151-2157.

USEPA. 2000. Methods for Measuring the Toxicity and Bioaccumulation of Sediment-associated Contaminants with Freshwater Invertebrates. EPA/600/R-99/064.

USEPA. 2003. Procedures for derivation of equilibrium partitioning sediment benchmarks (ESB)s) for the protection of benthic organisms :PAH mixtures. EPA-600-R-02-013.

USEPA. 2003. Guidance for developing ecological soil screening levels (ECO-SSLs). Attachment 4-1. OSWER Directive 9285.7-55. Washington, D.C.

USEPA. 2005. Ecological soil screening level (Eco-SSL) (various metals) guidance and documents. http://www.epa.gov/superfund/programs/risk/ecorisk/ecossl.htm

USGS. 1999. Areas susceptible to irrigation-induced selenium contamination in the western United States. US. Geological Survey, Crason City, Nevada..USGS Circular 1180.

Van der Hoeven, N. 1998. How to measure no effect. Part III: statistical aspects oh NOEC, EC_x and NEC estimates. Environmetrics 8: 255-261.

Willett, K., P. Gardinali, L. Lienesch, and R. DiGiulio. 2000. Comparative metabolism and excretion of benzo(a)pyrene in 2 species of ictalurid catfish. Toxicol. Sci. 58: 68-76.

NSPW Responses to Agency Comments
Baseline Ecological Risk Assessment
Draft Remedial Investigation Report
Ashland/NSP Lakefront Superfund Site

Tables

Table 1. Comparison of risk to *H. azteca* and *L. variegatus* Using Two Different Methodologies

Station	Matrix	*H. azteca* % Survival[1]	*L. variegatus* TLM HQ[2]
SQT6	Sand	91.3, 90 (90.7)	0.9-0.12
SQT2	Wood	91.3, 83.8 (87.6)	0.05-0.09
SQT3	Wood	72.5, 87.5 (80)	0.51-0.54
SQT4	Wood	95, 90 (92.5)	0.28-0.46
SQT5	Wood	91.3, 81.3, 81.3 (84.6)	1.14-1.59
SQT8	Wood	77.5, 85 (81.3)	0.08-0.1

1) Based on 28-d bioassays

2) Based on NEBR and Target Lipid Model (DiToro and McGrath 2000)

TABLE 2. *A Posteriori* Power Analysis of ANCOVA

Biological Metric	Variance Explained by Sample Station	Variance Explained by Station and NOCPAH	Actual Effect Size	Power to Detect a Small Effect	Variance Explained by Station and TotalPAH	Actual Effect Size
Density (estimated abundance)	0.69446	0.71771	0.02325	0.58	0.70295	0.00849
Number of Taxa	0.63395	0.63396	0.00001	0.5	0.63395	0
Number of ETO Taxa	0.5699	0.57244	0.00254	0.44	0.57034	0.00044
Number of ETO+Chironomidae Taxa	0.53295	0.54034	0.00739	0.41	0.53295	0
Percent Dominance	0.56409	0.5642	0.00011	0.44	0.56953	0.00544
Percent Chironomidae	0.8484	0.85441	0.00601	0.87	0.84845	5E-05
Percent Oligochaeta	0.69626	0.70025	0.00399	0.58	0.69725	0.00099
Percent Non-insects	0.84205	0.84876	0.00671	0.87	0.84229	0.00024
Number of Crustacea+Mollusca Taxa	0.47486	0.47491	0.00005	0.37	0.4774	0.00254
Percent Crustacea+Mollusca	0.86975	0.86979	0.00004	0.94	0.87389	0.00414
Shannon Diversity Index	0.67942	0.68258	0.00316	0.59	0.6805	0.00108
Shannon Equitability	0.58817	0.59444	0.00627	0.45	0.58966	0.00149
Biotic index (HBI)	0.63016	0.63064	0.00048	0.5	0.63165	0.00149
Percent Intolerant (<3.51)	0.57394	0.57427	0.00033	0.44	0.57408	0.00014
Percent Facultative (3.51-6.50)	0.80509	0.80595	0.00086	0.78	0.80627	0.00118
Percent Tolerant (>6.50)	0.80769	0.80898	0.00129	0.78	0.8091	0.00141
Percent Filterers	0.63268	0.6343	0.00162	0.5	0.64102	0.00834
Percent Gatherers	0.64641	0.65763	0.01122	0.51	0.6499	0.00349
Percent Parasites	0.68239	0.68357	0.00118	0.56	0.70259	0.0202
Percent Predators	0.63023	0.65637	0.02614	0.5	0.64187	0.01164
Percent Scrapers	0.64705	0.65067	0.00362	0.51	0.65462	0.00757
Percent Shredders	0.41942	0.42382	0.00440	0.34	0.4196	0.00018
	Averages:		0.00503	0.57		0.003661

Table 3. Ratio of EPA EMAP PAHs to URS measured PAHs

Station	Sum URS	Sum EMAP	Ratio
NSP-SE-SQT1-0605	333.38	365.94	1.10
NSP-SE-SQT5-0605	94.54	100.72	1.07
NSP-SE-SQT5-0605-AD	86.47	93.02	1.08
NSP-SE-SQT7-0605	22.94	23.57	1.03
NSP-SE-SQT7-0605 Dup	26.73	27.42	1.03
NSP-SE-SQT8-0605	110.73	111.40	1.01
NSP-SE-SQT8-0605-AD	69.75	70.79	1.01
NSP-SE-SQT2-0605	5.24	5.64	1.08
NSP-SE-SQT2-0605-AD	6.05	6.60	1.09
NSP-SE-SQT3-0605	37.08	39.86	1.07
NSP-SE-SQT3-0605-AD	49.33	53.59	1.09
NSP-SE-SQT4-0605	29.09	31.27	1.07
NSP-SE-SQT4-0605-AD	37.67	40.76	1.08
NSP-SE-SQT6-0605	10.00	11.51	1.15
NSP-SE-SQT6-0605-AD	12.51	14.28	1.14
NSP-SE-SQT6-0605-AD Dup	11.72	13.48	1.15
NSP-SE-SQT9-0605	0.32	0.34	1.08
NSP-SE-SQT10-0605	0.04	0.05	1.15
NSP-SE-SQT11-0605	0.83	0.97	1.16
NSP-SE-SQT11-0605-AD	61.98	63.57	1.03
NSP-SE-SQT12-0605	0.02	0.02	1.15
NSP-SE-SQT13-0905	0.03	0.03	1.09
NSP-SE-SQT14-0905	0.08	0.10	1.16
		Average =	1.09

Figure 1. Combined SEH and URS *H. azteca*

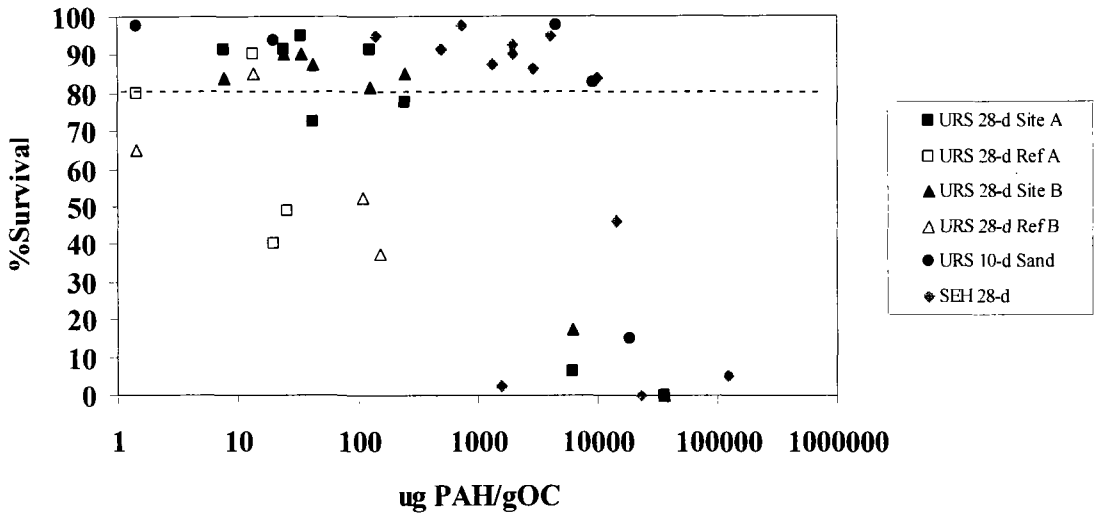

Figure 2. Combined SEH and URS *H. azteca* 28-d Wood

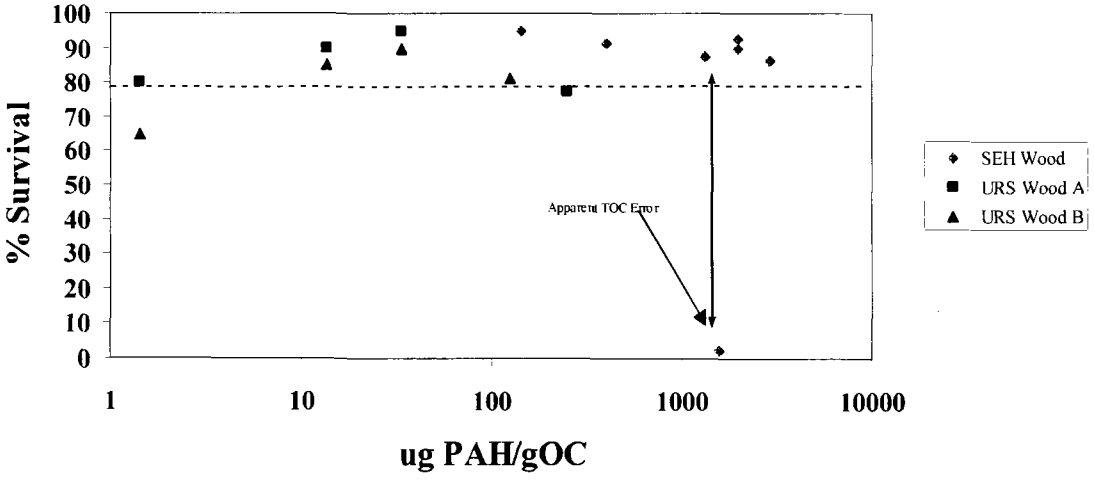

Figure 3. Combined SEH and URS *H. azteca 28-d* Sand, Normal Light

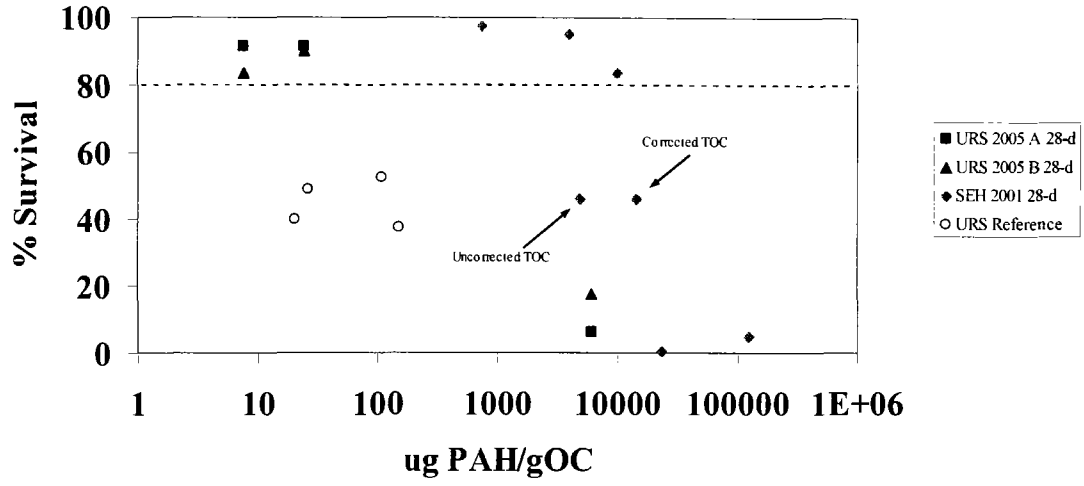

Figure 4. Combined 10-d and 28-d *H. azteca* Sand, Normal Light

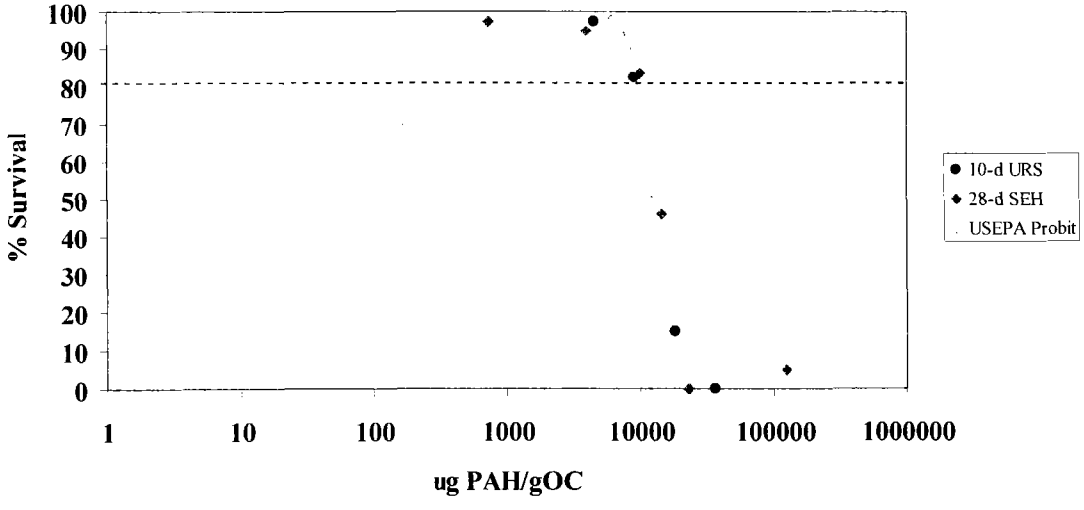

Figure 5. Comparison of PAHs in Sediment and *L. variegatus* lipids

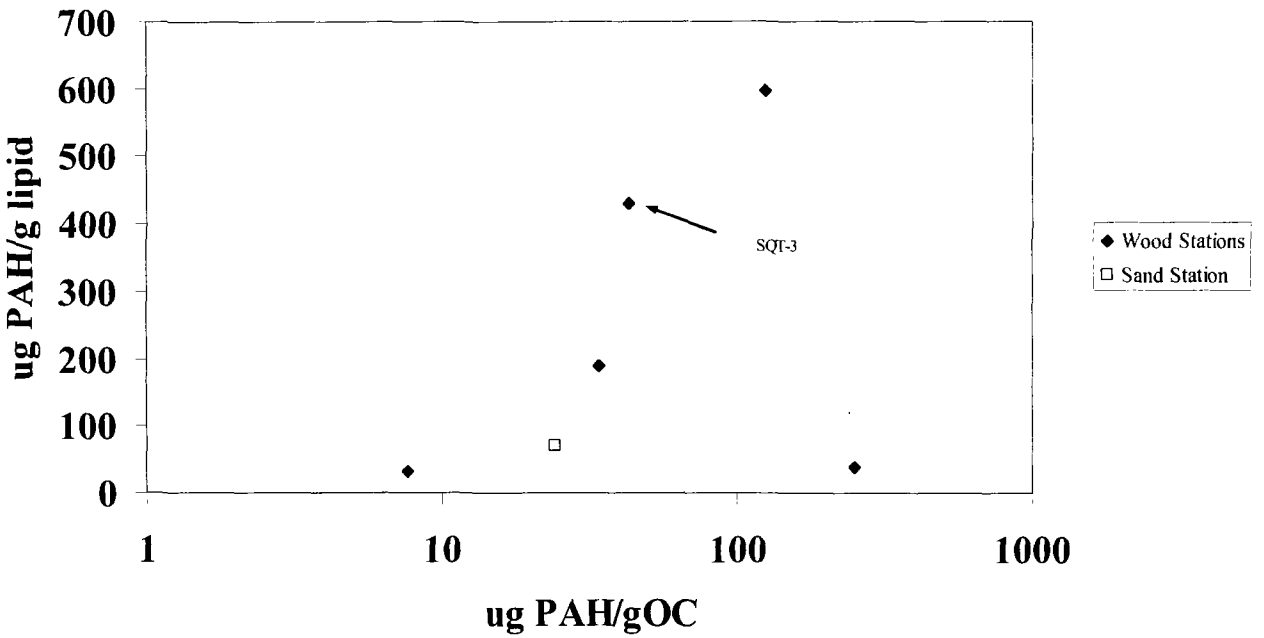

Figure 6 – Survival of *Hyalella* in 28-d Exposures Compared to EPA and URS Risk Benchmarks

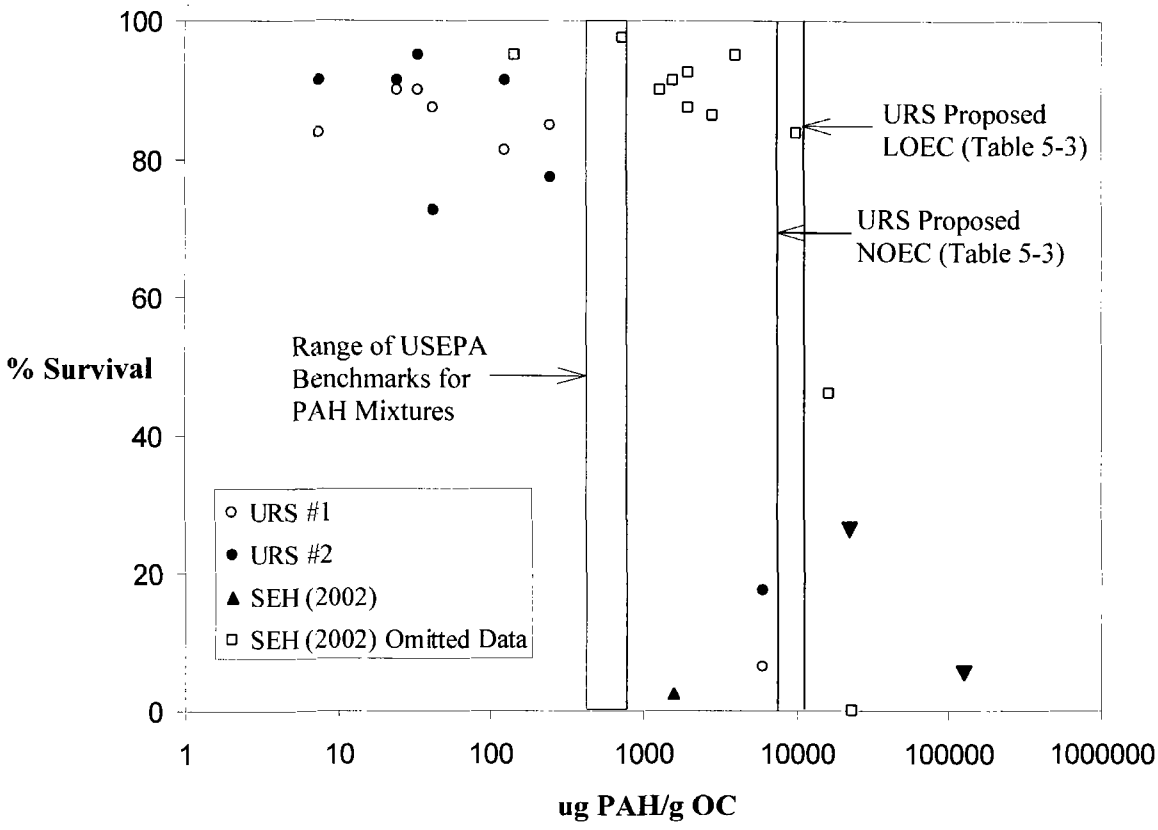

Figure 7. Combined SEH and URS *H. azteca* Sand, Normal and UV Light

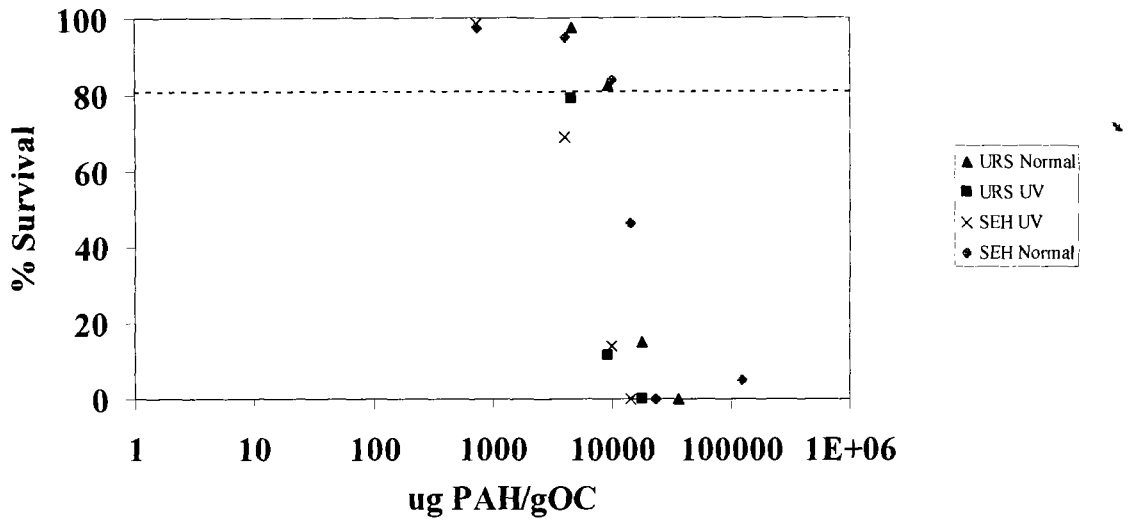

Figure 8. Effect of Refugia on UV Light-Induced Toxicity *H. azteca*